

**Andrés Enrique-Arias  
Manuel J. Gutiérrez  
Alazne Landa  
Francisco Ocampo**

***Perspectives  
in the Study of Spanish  
Language Variation***

***Papers in Honor of  
Carmen Silva-Corvalán***

***Verba***

***Anexo 72***

**2014**

# **Perspectives in the Study of Spanish Language Variation**

**Papers in Honor of Carmen Silva-Corvalán**

Edited by

ANDRÉS ENRIQUE-ARIAS  
*Universitat de les Illes Balears*

MANUEL J. GUTIÉRREZ  
*University of Houston*

ALAZNE LANDA  
*Universidad del País Vasco / Euskal Herriko Unibertsitatea*

FRANCISCO OCAMPO  
*University of Minnesota*

2014

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Esta obra atópase baixo unha licenza internacional Creative Commons BY-NC-ND 4.0. Calquera forma de reprodución, distribución, comunicación pública ou transformación desta obra non incluída na licenza Creative Commons BY-NC-ND 4.0 só pode ser realizada coa autorización expresa dos titulares, salvo excepción prevista pola lei. Pode acceder Vde. ao texto completo da licenza nesta ligazón: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.gl>

© Universidade de Santiago de Compostela, 2014

**Maquetación**

Ildefonso Vidal Ocampo

**Edita**

Servizo de Publicacións  
e Intercambio Científico  
Campus Vida  
15782 Santiago de Compostela  
[www.usc.es/publicacions](http://www.usc.es/publicacions)

**DOI** <https://dx.doi.org/10.15304/va.2014.701>

**ISSN** 2341-1198

## TABLE OF CONTENTS

<b>Preface</b> .....	7
<b>Introduction</b> <i>Andrés Enrique-Arias, Manuel J. Gutiérrez, Alazne Landa, Francisco Ocampo</i> .....	9
<b>I. MORPHOSYNTACTIC VARIATION AND CHANGE</b>	
<b>On the Grammaticalization of <i>dizque</i></b> <i>Asier Alcázar</i> .....	19
<b><i>Decir</i> en español rioplatense: un análisis semanto-pragmático</b> <i>Alicia Ocampo</i> .....	43
<b>Excess, Deficiency, and Mean: Resumptive Pronouns, Prepositional Phrase Chopping, and Gaps in Spanish Oblique Relative Clauses</b> <i>Álvaro Cerrón-Palomino</i> .....	75
<b>Rare Phenomena and Representative Corpora</b> <i>Bernard Comrie</i> .....	101
<b>Subject Position in Spanish: A Study of Factor Interactions with Prototypical Verbs</b> <i>Roberto Mayoral Hernández</i> .....	111
<b>Uso del sujeto y del objeto directo en el habla de once ciudades hispanohablantes</b> <i>Mercedes Sedano and Paola Bentivoglio</i> .....	137
<b>El orden de palabras en cláusulas subordinadas relativas con un sujeto y un verbo transitivo</b> <i>Francisco Ocampo</i> .....	177
<b>Cláusulas flexionadas predicativas con verbos de percepción</b> <i>María José Rodríguez-Espiñeira</i> .....	201

<b>Sobre las formas en -ra en el español de Galicia</b> <i>Guillermo Rojo and Victoria Vázquez Rozas</i> .....	237
---	-----

<b>Efectos del contacto de lenguas en el castellano de Mallorca: una perspectiva histórica</b> <i>Andrés Enrique-Arias</i> .....	271
---	-----

## II. SPANISH IN THE UNITED STATES

<b>La subida de clíticos en tiempo real y aparente en el español de Houston</b> <i>Manuel J. Gutiérrez</i> .....	301
---	-----

<b>Lateralización de /r/ implosiva: la conciencia fonológica y sus manifestaciones en el español puertorriqueño oral y escrito</b> <i>Antonio Medina-Rivera</i> .....	321
--	-----

<b>El español de tierras altas y tierras bajas: sus reflejos en el español de Los Ángeles</b> <i>Claudia Parodi</i> .....	341
--	-----

<b>Remarks on Pronominal Perseveration and Functional Explanation</b> <i>Ricardo Otheguy</i> .....	373
---	-----

<b><i>Dude was figureando hard:</i> El cambio y la fusión de códigos en la obra de Junot Díaz</b> <i>Domnita Dumitrescu</i> .....	397
--	-----

## III. LANGUAGE ACQUISITION

<b>Heritage Language Learners: Incomplete Acquisition of Grammar in Early Childhood</b> <i>Jürgen M. Meisel</i> .....	435
--	-----

<b>The English Phonological Skills of Latino-Spanish-English Dual-Language Preschoolers Living in Los Angeles: Implications for Practice</b> <i>Simona Montanari, Kaveri Subrahmanyam, Marlene Zepeda and Carmen Rodríguez</i> .....	465
---	-----

**Aspectual Distinctions in Spanish:  
A Case Study of Transfer in Second-Language Learners**  
*Ana Sánchez-Muñoz* ..... 497

**The Effect of Acoustic Training on the Acquisition of Second Language Spanish  
Stops**  
*John Stevens* ..... 515

#### **IV. LANGUAGE IN SOCIETY**

**Textos fundacionales de la indignación:  
Aproximación desde el análisis del discurso**  
*Francisco Moreno Fernández* ..... 539

**Discurso y negociación de identidad:  
grupos originarios de México en Carolina del Norte**  
*Liliana Paredes and D. Kyle Danielson* ..... 563



## Preface

This volume of peer-reviewed studies on Spanish language variation has been prepared as a tribute to Carmen Silva-Corvalán on the occasion of her retirement at the University of Southern California in 2014. Former students, colleagues, and friends have joined forces to produce these studies, all of them carried out within some of the research domains to which she has so fruitfully contributed. With this volume we intend to pay homage to someone who has been decisive in the training of many professionals in the field of Spanish linguistics.

The editors would like to thank Juan Luis Blanco Valdés and María José Rodríguez Espiñeira from the University of Santiago de Compostela for their invitation to publish this volume in the *Anexos de Verba* series. We are also grateful to Natalia Villar Conde and Ildefonso Vidal Ocampo for their assistance in the preparation of the manuscript for publication. Likewise we extend our gratitude to the colleagues that have refereed these papers and contributed their insights towards their improvement. And we also thank the authors for their enthusiasm and their patience with the different phases of the editorial project.

Querida Carmen: queremos agradecerte tu dedicación como profesora y tu cariño como colega y amiga. Estamos orgullosos de haber sido parte del proceso de elaboración de este volumen de homenaje a alguien tan importante para la lingüística hispánica y para nosotros, pero sobre todo lo estamos de haberte tenido en nuestras vidas. Te deseamos lo mejor en esta nueva etapa que comienzas.

*Andrés, Manuel, Alazne y Francisco*





## Introduction

Carmen Silva-Corvalán stands out as a prominent linguist among a select group of researchers in the humanities and the social sciences. Her sundry contributions to morphosyntax, pragmatics, and sociolinguistics, always based on solid empirical data collected in sociolinguistic interviews, have been highly influential. The methodologies that she has designed for the study of linguistic phenomena according to the most relevant social and linguistic variables have become models in the field. Moreover, she is a gifted teacher who has trained a significant number of leading scholars, all of them deeply influenced by her ideas, which they pass on to their own students.

It is difficult to select from her prolific output her most significant contributions as a linguist. As a whole, her collected studies have shed light on the analysis of linguistic simplification, transfer, grammatical convergence, the role of external and internal factors in processes of linguistic change, and other such phenomena.

One of Silva-Corvalán's main undertakings was the ELA (East Los Angeles) project, an endeavor funded by the National Science Foundation. Its main goal was to study processes of language change that were taking place in the Spanish spoken in Los Angeles by speakers of Mexican origin. As part of this research, Silva-Corvalán proposed a new definition of the variable *generation* that took into account the place where the speakers were born, the age of arrival to the United States in the case of immigrants, and the number of years in contact with English. Her working definition—it should be noted—is still used, with slight modifications in some cases, in the scholarship devoted to situations of language contact.

Silva-Corvalán studied for several years changes in the expression of verbal tense, mood and aspect across different generations of Spanish-English bilingual speakers from Los Angeles. Following a complex but

clear and detailed methodology in her analysis, the author uncovered important processes of simplification and loss in the verbal system of the community she was studying. She not only proved that the verbal system progressively weakened from one generation to the next, especially among US born bilinguals, but she also showed how the changes were taking place and revealed the strategies that speakers use to lighten the cognitive load generated by the use of two languages. This investigation provided solid evidence in support of the sociolinguistic claim that linguistic change can be observed.

Silva-Corvalán's studies about the *ser/estar* opposition are also well-known. With great linguistic perception, she noticed a process of semantic change in which speakers generalized the *estar* form so that it became part of the semantic domain traditionally covered by the *ser* form. She coined the term *estar innovador* (*innovative 'estar'*) to describe this new manifestation of the verbal copula and demonstrated that it was currently undergoing a process of linguistic change. Her pioneering publications in *Language* and other journals inspired subsequent investigations on the subject.

With her research on linguistic variation at the morphosyntactic level, Silva-Corvalán became an active participant in the sociolinguistics debate on the feasibility of extending variationist studies beyond the phonological component. Was it possible to apply Labov's model to variables in which the synonymy of the variants was questionable? Silva-Corvalán's studies on the conditional, the imperfect and other verbal forms in the various dialects of the Spanish-speaking world demonstrated that the variationist approach could be adapted to the nature of the new linguistic variables to be studied. To account for the complexity of morphosyntactic variables, she also deployed the distinction between referential meaning and contextual meaning: the latter is studied qualitatively to determine how it is affected by semantic-discursive and pragmatic factors. The problem of synonymy, or lack thereof, between variants is thus treated with approaches that incorporate pragmatic interpretations and the need to consider the discourse as object of analysis.

In more recent years, Silva-Corvalán has devoted the lion's share of her efforts to the study of language acquisition by children growing up in a two language environment. Having patiently collected the relevant linguistic data in the course of many years, her new book *Bilingual Language Acquisition: Spanish and English in the First Six Years* (Cambridge University Press 2014) now presents a detailed discussion of all phenomena related to the simultaneous acquisition of two linguistic systems.

There is no doubt that the main issue investigated by Silva-Corvalán has been that of linguistic change; her contribution to this field is unique. Driven by the desire to explain the motivation that lies behind linguistic change, her studies have followed a clear methodology that considers linguistic differences according to social and stylistic factors, providing concrete evidence on linguistic variation. The use of apparent time as an analytical tool has also allowed her to identify phenomena of linguistic change. She has devoted as well particular attention to situations of bilingualism, which help to understand the role played by the contact language, sometimes as a direct motivator and sometimes, when the change has an internal motivation, as a decisive factor in the speed at which the change is propagated. Her pondered conclusion—i.e. that transfer between languages is more likely to occur when the structures in the two languages are more similar—remains a lasting contribution to our understanding of this phenomenon.

The type of research developed, then and now, by Silva-Corvalán not only sheds light on how the grammatical system functions and the nature of the relationships between language and society, but also constitutes an important contribution to other disciplines in the social and natural sciences, and in the humanities in general. It has contributed, in the process, to developing new pedagogical methods for teaching Spanish as a second language and as a second dialect in the case of heritage students.

The collection of peer-reviewed studies contained in this volume aims to offer a few glimpses into the vast research area of Spanish language variation which she has so fruitfully explored in one way or another. The contributors to the volume and the editors have maintained a lasting professional dialogue with Carmen Silva-Corvalán; indeed, an important number of contributors in this volume were also her students.

The material has been organized into different sections although it is hard to draw neat categories as some papers could well be in more than one section. The four areas are (I) Morphosyntactic Variation and Change, (II) Spanish in the United States, (III) Language Acquisition, and (IV) Language in Society.

The opening section *Morphosyntactic Variation and Change* contains ten studies. In “On the Grammaticalization of *dizque*,” Asier Alcázar presents a twofold analysis of this Latin American evidential expression: on the one hand, as a particle and, on the other, as the exponent of a clause-level projection. At the same time, the author reexamines this locution’s origin, evolution and grammaticalization process in several varieties. Alicia Ocampo

also analyzes a reflex of the Spanish verb *decir* ‘to say’ in her paper «*Decir en español rioplatense: un análisis semanto-pragmático*». In her analysis *decir* is seen as an evidential device used in most categories described within a taxonomy of evidentiality considered in its broad sense. The author finds that the form *dice* is undergoing a number of changes towards discoursivization, sharing properties with other markers that followed the same diachronic path. This leads her to conclude that, in the variety spoken in the area of Buenos Aires, this verbal form is becoming a particle with a hearsay index function.

The next two papers deal with classical problems in studies of Spanish syntactic variation: resumptive pronouns in relative clauses, and the variation in the use of the “personal *a*”. In “Excess, Deficiency, and Mean: Resumptive Pronouns, Prepositional Phrase Chopping, and Gaps in Spanish Oblique Relative Clauses,” Álvaro Cerrón-Palomino considers three alternations found in Spanish oblique relative clauses: (i) the presence of resumptive pronouns, which are mainly favored by animate antecedents (ii) gap-containing relative clauses, favored by constructions containing only verbs, and (iii) relative clauses with prepositional phrase chopping, favored by inanimate antecedents. On the other hand Bernard Comrie’s contribution, “Rare Phenomena and Representative Corpora,” deals with the methodological difficulties in corpus-based studies of “personal *a*” in Spanish ditransitive constructions with a human theme because the construction is rare and uneven in texts. The reasons for this difficulty are set out, as well as the advances that would be needed in order to overcome them.

The next three chapters deal with variation in the order of different syntactic constituents, namely subject and object in a number of syntactic environments. In “Subject Position in Spanish: A Study of Factor Interactions with Prototypical Verbs” Roberto Mayoral makes use of an interdisciplinary approach to research into the impact of syntactic, semantic and contextual factors on subject position in Spanish; what he finds is that all of them –with the exception of the animacy of the subject referent– are statistically significant. In their article “Uso del sujeto y del objeto directo en el habla de once ciudades hispanohablantes,” Mercedes Sedano and Paola Bentivoglio analyze the position of subject and direct object with respect to the verb in different Spanish varieties, and they find numerous similarities that seem to indicate a deeply-rooted homogeneity in the language in spite of dialectal variation. In “El orden de palabras en cláusulas subordinadas relativas con un sujeto y un verbo transitivo,” Francisco Ocampo presents a motivation for

word order variation in relative clauses occurring in a corpus of Rioplatense Spanish informal conversations. The preverbal position of the subject is motivated by its topicality. There are also two other factors that correlate with the preverbal position of the subject, related to the fact that non-restrictive relative clauses have greater syntactic autonomy.

In the next article, “Cláusulas flexionadas predicativas con verbos de percepción,” María José Rodríguez-Espiñeira examines the sentence-type *V NP que S*, where the lexical NP is the direct object of the main verb and the antecedent of the relative *que*, which is the subject of the relative clause. Based on the CREA corpus, the author proposes that this construction is pragmatically motivated, accomplishing a presentational function in discourse.

The last two chapters in this section deal with contact varieties of Peninsular Spanish. Guillermo Rojo and Victoria Vázquez Rozas, in “Sobre las formas en *-ra* en el español de Galicia,” analyze the uses of Spanish *cantara* in relation to both *cantase* and *había cantado* type forms in the Spanish in contact with Galician in northeastern Spain, using sociolinguistic interviews with upper middle-class speakers. In the last contribution of this section, the study of Andrés Enrique-Arias, “Efectos del contacto de lenguas en el castellano de Mallorca: una perspectiva histórica,” uses several phenomena of the variety of Spanish in contact with Catalan in Majorca (the verb *pedir* ‘ask for something’ with the semantic and syntactic structure of *preguntar* ‘ask a question’ and the preponderance of synthetic future *cantaré* over periphrastic *voy a cantar*) as case studies in order to emphasize the relevance of a historical perspective in the study of contact varieties.

The second section of this collection, *Spanish in the United States*, includes five articles that have studied oral or written samples of Spanish from several cities in this country (Los Angeles, New York, and Houston). In “La subida de clíticos en tiempo real y aparente en el español de Houston,” Manuel J. Gutiérrez analyzes the position of verbal clitic pronouns in verb phrases that allow clitic climbing in data drawn from conversations with bilingual speakers in Houston (two intergenerational samples separated by a few decades). A quantitative analysis of the data allows the author to identify the current tendencies that motivate the speakers’ selection and the changes that are taking place in the position of clitic pronouns. Antonio Medina-Rivera studies the lateralization of /r/ in his article “Lateralización de /r/ implosiva: la conciencia fonológica y sus manifestaciones en el español puertorriqueño oral y escrito.” The author finds that, in the case of speakers

of Puerto Rican origin residing in Cleveland, the lateralization goes beyond the oral component because written samples selected from social networks indicate that the phenomenon may represent a feature of national identity. In “El español de tierras altas y tierras bajas: sus reflejos en el español de Los Ángeles,” Claudia Parodi examines data produced by Salvadorian and Guatemalan speakers as well as speakers of Mexican origin to study the contact situation between two Spanish dialects (lowlands and highlands) in Los Angeles. The author observes that some of the linguistic features from the lowlands (Central America speakers) are changing in the new generations due to contact with highland speakers (Mexican origin), the dominant Spanish variety in the area. Ricardo Otheguy studies the presence/absence of subject pronouns that are placed preposed or postposed to a Spanish finite verb in data from Spanish speakers residing in New York City in “Remarks on Pronominal Perseveration and Functional Explanation.” According to the author, the general results of his study support the perseveration hypothesis in some structures (V V), but not in the case of [pV pV] series, which allows the investigator to explore a number of variables that may explain the motivations of speakers when producing sequences of these types. In the last article in this section “*Dude was figureando hard: el cambio y la fusión de códigos en la obra de Junot Díaz*,” Domnita Dumitrescu examines the way this Dominican writer makes use of Spanish in several of his English-language works. Dumitrescu concludes that Junot Díaz’s writing is a very good example of translanguaging, that is, the use of a number of discourse techniques that cannot be linked to any specific language in a bilingual context, but are rather typical of the linguistic production of the bilingual speakers themselves.

The third section of this volume, *Language Acquisition*, incorporates four articles that tackle relevant issues concerning multilingual situations and classroom environments. In “Heritage Language Learners: Incomplete Acquisition of Grammar in Early Childhood,” Jürgen Meisel offers an in-depth discussion of the acquisition problems observed in speakers who have grown in environments where two languages are spoken, one at home and another by the general community. The author revisits the previously-held claim that heritage speakers show incomplete acquisition and a simplified system due to insufficient exposure to the target language and shows that the notion of incomplete or reduced system has not been clearly defined in previous research. The phonological skills of bilingual Los Angeles preschoolers are examined by Simona Montanari, Kaveri Subrahmanyam,

Marlene Zepeda and Carmen Rodríguez in “The English Phonological Skills of Latino-Spanish-English Dual-Language Preschoolers Living in Los Angeles: Implications for Practice.” The researchers find that during language development in bilinguals, Spanish and English phonological systems interact, facilitating some phonological skills, while simultaneously delaying others. The article by Ana Sánchez-Muñoz, “Aspectual Distinctions in Spanish: A Case Study of Transfer in Second-Language Learners,” presents a study of the acquisition of the progressive constructions and the imperfect/preterite opposition by students of Spanish as a second language. The author tries to determine how the process of transfer from English to the target language (Spanish) works in this aspect of the grammar. Her study suggests that the outcome of transfer depends on the similarities of the structures considered in the two languages in contact. John Stevens’ study, “The Effect of Acoustic Training on the Acquisition of Second Language Spanish Stops,” examines whether speech analysis computer software leads to gains in the pronunciation of second language Spanish stop consonants of English-speaking learners by focusing their attention on the differences that exist between English and Spanish stops in terms of voice onset time (VOT). The results provide some justification for the inclusion of acoustic training in college-level Spanish phonetics courses that aim to improve student L2 Spanish pronunciation.

The final section of this volume, *Language in Society*, includes two studies that analyze discourse produced by specific groups of speakers in their interactions with society. Francisco Moreno Fernández in his article “Textos fundacionales de la indignación: Aproximación desde el análisis del discurso” studies the universe of ideas and actions related to the social movement known as *movimiento de los indignados*, *Spanish revolution* or *#15M* with special attention to those associated with language and communication in order to characterize *indignant discourse*. In the last work of this collection, “Discurso y negociación de identidad: grupos originarios de México en Carolina del Norte,” Liliana Paredes and D. Kyle Danielson base their study on tape-recorded conversations with Mexican immigrants whose first tongue is an Amerindian language and for which Spanish constitutes a second language in a contact situation with English. The authors underscore the complexity of this situation in analyzing how these speakers’ discourse manifest their cultural and ethnic differences and the way they negotiate their identity in order to gain access to services in conditions similar to those of other immigrants.



In sum, this collection of studies showcases just a few of the different facets comprising nowadays the study of Spanish language variation. Much of the research presented in here owes something to the work of Carmen Silva-Corvalán. It all goes to show how some of the most notable contributions to the field in the last decades have been inspired by the scholarly efforts of our distinguished mentor.

*The Editors*

# I

## **Morphosyntactic Variation and Change**



## On the Grammaticalization of *dizque*

Asier Alcázar  
*University of Missouri*

**Abstract:** This paper studies the grammaticalization of Latin American Spanish *dizque* ('they say that/it is said that', a quotative/reportative strategy) with particular reference to data from secondary sources in Bolivian, Colombian, Ecuadorian and Mexican Spanish, as well as diachronic data from Old Castilian. The evidential and epistemic properties of *dizque* are discussed in relation to current theoretical debates regarding whether a continuum exists between lexical and grammatical evidentiality. An alternative analysis of *dizque* is presented as evidence for such a continuum in contrast to that of Aikhenvald (2007), who analyzes *dizque* as an emergent evidential that may not eventually become a grammatical evidential. By contrast, the analysis defended in the current paper proposes that, while *dizque* may be an emergent evidential in some dialects, in others it seems to have become a grammatical evidential whose expression is not (yet) obligatory.

**Keywords:** *dizque*, evidentiality, grammaticalization, Spanish.



## 1. INTRODUCTION

This introduction discusses the analysis of *dizque* as an emergent evidential in the larger context of competing definitions of evidentiality.

In its *narrow definition* evidentiality refers to the source of information that a speaker has for her utterance (Aikhenvald 2004: 1). While speakers have access to new information from a wide array of socially and culturally relevant sources (such as parents, teachers or tribe leaders), the recorded grammaticalized sources of information are few. According to Aikhenvald (2007: 4), grammaticalized sources comprise information acquired through the senses, inferences and assumptions, information relayed by a second/third person or hearsay (reportatives), or information reproduced verbatim (quotatives). Less studied forms of evidentiality extend to actions the speaker is carrying out or just carried out at utterance time (performatives, Oswalt 1986), and to general or traditional knowledge (Givón 1982, Anderson 1986, Lazard 2001, Faller 2002). All of the above sources can be collapsed under the broader labels direct vs. indirect evidence, as in Willet's taxonomy (1988). Systems with two evidential choices encapsulate these two categories in Aikhenvald's survey (2004: 154-59).

In its *broader definition* (Chafe 1986: 262, Mithun 1986: 89, Friedman 1986: 168-9), evidentiality is extended to include speaker attitude and stance, such as commitment to the veracity of information, an assessment of the reliability of the information or the source, and face-saving strategies (Brown and Levinson 1987; such as displaced responsibility or hedging). The broader definition opens the door to consider other evidential expressions where evidential meaning may not be primary. At the same time, the core meaning of evidentiality is less restricted, undermining the study of evidentiality as an independent grammatical category or well-defined morphosyntactic category.

Aikhenvald attributes the broader definition not to grammatical evidentials, but to other evidential-like expressions instead, which do not display a primary or exclusive evidential meaning. Aikhenvald discusses *emergent evidentials* and *evidentiality strategies*, forms that do not conform to the narrow definition (2004, Ch. 4). For Aikhenvald forms such as Mexican and Colombian *dizque* 'they.say.that' (quotative, reportative cf. Travis 2006, Olbertz 2007), where evidential meaning may become primary, are *emergent evidentials*, or *evidentials in the making* (pp. 140-2). Regarding the semantics of *dizque*, Aikhenvald notes that 'In both varieties *dizque* extends to cover doubt and (negative) attitude to the information and its

validity—that is, it goes beyond simply reporting what someone else has said’ (2007: 14). Because of their broader interpretation, emergent evidentials do not conform to the narrow definition.

Regarding evidentiality strategies, Aikhenvald notes that other grammatical categories such as tense, non-indicative mood, person, as well as verbs, nominalizations or passives have secondary evidential meanings or pragmatic extensions. For example, the conditional in French can denote indirect evidence and uncertainty about the reliability of the information. In German it can be interpreted as a reportative. The future in Spanish can express inference and uncertainty (pp. 107-9).

Evidentiality strategies and emergent evidentials fall into the broader category of lexical evidentiality, together with the recruitment of parts of speech to express evidential meanings (Chafe 1986, Biber and Finegan 1989, Nuyts 2001, Squartini 2007a, Diewald and Smirnova 2010). Broadly defined, verbs (*I hear, it seems, they say...*), modal verbs (*must, can, may...*), adverbials (*apparently, purportedly, on all accounts...*), modal adverbs (*probably, possibly, surely...*), adjectives (*apparent, alleged, reportedly...*) and modal adjectives (*possible, probable, sure...*) convey a set of meanings to include evidentiality, probability, speaker attitude and stance. “A sort of implicit confrontation” (Squartini 2007b: 1) exists regarding whether lexical evidentiality is to be studied within the umbrella of (grammatical) evidentiality (see Aikhenvald 2007 for a dissenting view).

This paper studies the process of grammaticalization of *dizque* in Latin American Spanish. It compares Mexican and Colombian Spanish with Bolivian and Ecuadorian Spanish. It proposes that, in some dialects, the semantics of *dizque* and its syntactic distribution is much closer to a grammatical evidential than an emergent evidential. *Dizque* may be the exponent of a particle-like category in some cases, with flexible syntax and variable scope, but a sentential category in other cases; the latter case being virtually synonymous with a grammatical evidential. This analysis, in turn, begs questions about whether certain characteristics of emergent evidentials that set them apart from grammatical evidentials could be part of a natural path to the grammaticalization of evidentiality as a sentential category.

The outline of the paper is as follows. Section 2 discusses the grammaticalization, meaning and origin of *dizque*, with particular reference to Bolivian, Colombian, Ecuadorian, Mexican and Peninsular Spanish, but also Latin American dialects more generally. Section 3 presents a discussion of the data and an alternative analysis to Latin American *dizque* being not

exclusively an emergent evidential. Section 4 presents the conclusions of the paper and directions for further research.

## 2. LATIN AMERICAN *DIZQUE*: GRAMMATICALIZATION, MEANING AND ORIGIN

This section begins discussing whether *dizque*, on account of its phonological reduction, is a verb of saying adopting a grammatical function (§2.1). Recent literature proposes an evidential and/or epistemic function. Although *dizque* is a good fit as an emergent evidential, in some dialects this analysis leaves some characteristics unexplained. The semantics and syntactic distribution of *dizque* is presented first in Colombian, Mexican and Bolivian Spanish (§2.2); then in Ecuadorian Spanish (§2.3). The grammaticalization of *dizque* begs the question of whether its origin is a native development or a borrowing (§2.4). The answer to this question is critical in the characterization of *dizque* as an emergent vs. grammatical evidential, as some of these languages, such as Quechua, have been described to have grammatical evidentiality. The last section ends with a summary (§2.5).

### 2.1. *Dizque* < dice(n) que: A process of grammaticalization?

*Dizque* appears to be a reduction of present tense forms of *decir* ‘say’, either in the third person (1a, 1b) or in the short passive<sup>1</sup> (1c).

- (1) a. *Dic-e que*  
 say-THV COMP  
 ‘She/he says that’  
 b. *Dic-e-n que*  
 say-THV-3PL COMP  
 ‘They say that’  
 c. *Se dic-e que*  
 3SG.IMP say-THV COMP  
 ‘It is said that’

The phonological reduction of (1) into *dizque* may be taken as a sign that a process of grammaticalization is at play (Hopper 1991, Traugott and Dasher 2002, Bybee 2003). Phonological reduction can be understood as

<sup>1</sup> Spanish *pasiva refleja*. An anonymous reviewer notes that he or she is more familiar with the translation *passive-reflexive*.



an epiphenomenon of a rise in frequency of use. When words, phrases and expressions are used often, this usually leads to losses in prosodic stress and articulatory gestures. For example, in the grammaticalization of the English future auxiliary *going to* from the verb of motion *go*, this form is reduced to *gonna*. In this particular case, phonological reduction is caused by the adoption of a new grammatical function: expression of future tense.

Regarding phonological reduction, other attestations of *dizque* include *diz*, *dice*, *quesque*, and *desque*; also *disk*, *diske* in the new registers of the Internet. Because *dizque* is the most frequent variant, and it is seemingly attested across different dialects, it has become the emblematic form or the citation form. However, note that *dizque* does not constitute the last possible stage in phonological reduction. Phonological erosion continues with forms such as *izque*, *ihque* or *i que*, which were attested as early as the mid 20<sup>th</sup> c. (Kany 1944).

Phonological reduction does not always entail grammaticalization. This is the case of the English noun *Tallica* (from the band *Metallica*), or the expression *That is a girl!* (used in Hitchcock's *The Birds*, 1963) vs. present-day English *Atta girl!*. Similar examples of phonological reduction for Spanish, which would not signal grammaticalization (or reanalysis), include present indicative forms of the verb *estar* 'be (stage-level)'—*toy* for *estoy* 'I am', *ta(s)* for *estás* 'you are', *tamo(s)* for *estamos* 'we are'—or expressions such as *pa'na'* (*para nada* 'lit. for nothing') 'at all'.

If phonological reduction signals that *dizque* is participating in a process of grammaticalization, the grammatical function that *dizque* expresses would need to be clearly identified. Recent literature has taken note of *dizque* in the wake of a surge of research on evidentiality as a grammatical category (Aikhenvald 2004). According to recent evidential literature (Travis 2006, Olbertz 2005, 2007, Babel 2005, Lamy 2008, inter alios), *dizque* expresses evidential (reportative, quotative) and/or epistemic meaning (speaker attitude/stance, probability, certainty). Accordingly, phonological reduction of (1) may be a sign of a process of grammaticalization toward a (pseudo-) grammatical exponent of evidentiality and/or epistemic modality.

## **2.2. On the semantics and syntactic distribution of *dizque*: Colombian, Mexican, and Bolivian Spanish**

Regarding the semantics of *dizque*, and focusing on its evidential import for now (vs. epistemic meaning), recent evidential literature consistently identifies *dizque* as a quotative (direct/indirect speech) and a reportative

marker. Some scholars explicitly note the absence of non-reportative/non-quotative interpretations. For example, Babel (2009) reports that ‘inference and speculation by the speaker never appear with reportative *diz(que)*’ (p. 16). Cruschina and Remberger (2008) make a similar case for Romance equivalents of *dizque*. But, as we will see in the current section and the next, there is much diversity in the syntax and semantics of *dizque*. Aikhenvald (2007) observes that *dizque* can express inference and assumption in some dialects. She views *dizque* more generally as a non-direct evidence marker. *Dizque* also appears to have mirative uses that have not yet attracted much attention (see Kany 1944).

With regard to its syntactic distribution, generalizing over different dialects, *dizque* tends to appear preverbally. This was also the case in Old Spanish (Olbertz 2007). When *dizque* is understood to modify a clause, it may appear before or after the clause it modifies; before or after the verb; or inserted as a parenthetical. On the other hand, when *dizque* is interpreted to modify phrases or heads, it typically precedes or follows them. The syntactic distribution of *dizque* is diverse, among other reasons, because *dizque* may be able to modify heads, phrases or clauses. But it is important to stress that its syntactic distribution is dialect-specific. The current section and the next will focus on some of these dialectal differences (for further details, the reader is referred to the cited literature).

The evidential interpretations of *dizque* for Colombian, Mexican and Bolivian Spanish are illustrated next. In these dialects, evidential meaning presents with or without a simultaneous epistemic reading. The initial focus will be on examples where *dizque* is understood to modify sentences. In these cases, the evidential reading seems to be primary (over epistemic meaning) or exclusively evidential (even if pragmatically compatible with an epistemic interpretation in some contexts). However, it is not always clear what authors mean by the meaning of *dizque* when the compatibility of interpretations is not explicitly confined to semantics vs. pragmatics, or when the authors do not (explicitly) indicate whether semantics and pragmatics are independent, separable domains in the theoretical analysis they entertain. In what follows, we need to bear this caveat in mind.

In Colombian Spanish, *dizque* introduces direct or indirect speech and reported information (Travis 2006). An example of the quotative function is given in (2); hearsay is exemplified in (3).

- (2) A: *Y yo dizque*  
 M: @@@  
 A: [XXX], <VOX *mi amor, A qué horas fue que llegamos* VOX:??>  
 Y <@ *no, dizque @*>, <VOX *No=, hace como dos o tres*  
*horas VOX*>  
*Y hacia como media hora acabábamos de llegar.*  
 Angela: ‘And I *dizque*,’  
 María: @@@  
 Angela: ‘XXX, “Darling, what time did we get home?” And  
*no, dizque*, “No, two or three hours ago.” And we’d only  
 just got home about half an hour before.’  
 Travis (2006: 1281-2; ex. 6)
- (3) *porque dizque iba a enterrar a una persona*  
 ‘because *dizque* she was going to bury a person’  
 Travis (2006: 1282; ex. 9)

Travis makes the following observation concerning the potential for *dizque* to be interpreted epistemically: ‘as with *dizque* with reported speech, it [hearsay *dizque*] can mark material that the speaker believes to be true as well as material that is doubted’ (p. 1284).

Working on Mexican Spanish, Olbertz (2007) focuses on the reportative function of *dizque* (4, 5), but she also produces some examples of indirect speech (6).

- (4) [at work, a nanny talks about her own children]  
*Siempre tuvieron celos, dizque más me ocupaba de éstos*  
*que de ellos.*  
 always had jealousy more 1SG.REFL care.1SG.PST of these than of  
 them  
 ‘They have always been jealous, *saying* I cared more for these  
 children than for them.’  
 Olbertz (2007: 155, ex. 6; cf. *Fotografía en la playa* by Emilio  
 Carballido)
- (5) *Los terrenos dizque eran de la esposa del superministro*  
*Ulises López.*  
 The land were of the wife of.the superminister  
 Ulises López  
 ‘The land was *said to* belong to superminister Ulises López’s  
 wife.’



with the verb *decir* ‘say’ in quotative uses; and a higher frequency of disclosure of source of information, which may indicate *dizque* is becoming obligatory. Another difference is that, for Bolivian, there appear to be no tendencies to interpret heads and phrases epistemically and clauses evidentially (Babel p.c.). These are discussed next for Colombian and Mexican Spanish.

Travis (2006) comes across certain uses of *dizque* in Colombian that she refers to as ‘labeling’. Instead of modifying a clause, and serving a quotative or reportative function, the evidential expression is in some cases understood to modify a phrase, and is interpreted epistemically.

In what I have termed “labeling” *dizque* does not introduce a full clause (as is its most common use for reported speech and hearsay), but a noun phrase or a prepositional phrase. [...] *Dizque* in this use also differs from that of marking reported speech and hearsay in that it does not directly attribute the label to a (specific or nonspecific) third party, but simply indicates that it is not attributable to the speaker. Thus, *dizque* is not used with labels that the speaker is prepared to vouch for, and the notion of doubt has been conventionalized, in some cases extending to falsity. [...] It encodes something like “other people say this; I don’t want to say: I know this”, and is similar to English “so-called”. (Travis 2006: 1287)

Here are some of her examples illustrating *labeling* uses (8-10 cf. Travis 2006: 1287-9; exs. 18, 23, 24).

- (8) ... *en medio de la confusión y de la angustia, caí en manos de una mujer médica bioenergética que a la vez era dizque sofróloga.*  
 ‘... in the confusion and distress I fell into the hands of a female bio-energy doctor who was also a *dizque* sophrologist.’  
 (Castro Caycedo 1994: 236)
- (9) A: ... *Se presentó como, dizque narcotraficante, Yo me llamo no sé qué, narcotraficante,*  
 Angela: ‘... He introduced himself as *dizque* a drug dealer.  
 I’m called so-and-so, drug dealer,’  
 (Colombo: 469–472)
- (10) ... *se abrió dizque a comerciar. Comerciar era llegar a los almacenes, pedir, no pagar, y exigir dinero de vuelta... les sacaban dinero dizque de vuelta.*

‘... she began *dizque* to do business. Do business was to go into the shops, make an order, not pay, and demand change... they would get money out of them *dizque* in change.’  
(Castro Caycedo 1994: 100–101)

For Mexican Spanish, Olbertz reports and interplay between the interpretation and the size of the linguistic expression *dizque* is understood to modify. The epistemic interpretations become more prominent as the size of the expression decreases. Parallel to that, the evidential interpretations become secondary or even excluded.

[...] we can conclude that as the size of the linguistic expression modified by *dizque* decreases, the subjective relation between the communicated content and the speaker (i.e. the parameter ‘speaker attitude’) becomes gradually more prominent, while the objective relation (i.e. the two parameters referring to information source) becomes increasingly irrelevant. More specifically: both the objective and subjective relations play a role at the level of the main clause; the subjective relation becomes more prominent when *dizque* modifies subordinate clauses and constituents, and it becomes the central semantic value in the case of the modification of predicates [heads]. (Olbertz 2007: 165–6)

Regarding the modification of heads, Olbertz reports a primary epistemic interpretation for adjectives (11) and an exclusively epistemic interpretation for verbs (12) and predicative adjuncts: ‘the expression of epistemic modality has in fact become the exclusive function of *dizque*’ (p. 161). This claim is similar to the one Travis makes for Colombian when *dizque* has phrasal scope.

(11) - *Pues, ¿qué oíste?*

*Una cosa que dijeron los del gobierno ese dizque provisional*

‘- So, what did you hear?’

Something that those people of that *supposedly* provisional government said’ Olbertz (2007: 161, ex. 22; cf. *La casta divina* by Felipe Victoria Zepada)

(12) *A los seis meses de andar dizque gobernando se puso enfermo.*

‘After having gone about *pretending* to rule for six months he fell ill.’

Olbertz (2007: 162, ex. 25; cf. *Arráncame la vida* by Ángeles Mastretta)

With phrases (13-4), Mexican *dizque* tends to be used as reported speech and hearsay. But it seems to carry a conversational implicature: ‘a negative truth commitment on the part of the speaker’ (p. 158).

(13) *Andrés estaba rojo dizque del coraje pero era del brandy.*

Andrés was red of.the anger but was of.the brandy  
‘Andrés had turned red *supposedly* from anger, but it was from brandy.’

Olbertz (2007: 158, ex. 17; cf. *Arráncame la vida* by Ángeles Mastretta)

(14) *Claro que no faltó quien intent-ara llev-a-r*

of.course that NEG lack.3SG.PST who attempt-SUBJ  
carry-THV-INF *pulque o mezcal dizque para el difunt-ito*  
pulque or mezcal for the deceased-DIM

‘Of course there were also people who would try to bring pulque or mescal, *supposedly* for the dear dead’

Olbertz (2007: 159, ex. 18; cf. *Excelsior*, “Día de Muertos en México”, Nov. 1, 1996)

The modification of heads and phrases makes Mexican *dizque* a poor fit as a grammatical evidential. Grammatical evidentials (narrowly defined) modify clauses (denote sentential scope). They do not allow phrasal scope (Alexandra Aikhenvald, p.c.). With particular reference to scope possibilities, Mexican *dizque* seems more directly comparable to adverbs, such as *aparentemente* ‘apparently’ or *supuestamente* ‘supposedly’.

### 2.3. On the semantics and syntactic distribution of *dizque*: Ecuadorian Spanish

Ecuadorian is different from the previous three dialects in that *dizque* cannot modify heads and phrases; it can only be understood to modify clauses (Olbertz 2005). Ecuadorian is atypical in that the position of *dizque* is virtually fixed preceding the verb, including negative sentences (15).

(15) *No dizque pod-ían pag-a-r a nadie*  
NEG can-3PL.PST pay-THV-INF to nobody

‘*They say* they could not pay anybody’

(Olbertz 2005: 7, ex. 27; spoken corpus cf. Muysken 1978)



Olbertz concludes that *dizque* + V must have been reinterpreted as a complex verbal construction, which explains the position of negation before the first element of this construction (p. 90; my translation<sup>2</sup>)

Similarly to Bolivian Spanish, in Ecuadorian Spanish the quotative use is gradually becoming an obligatory collocation with *decir* ‘say’ (16). In Olbertz’s corpus sample *dizque* appears by itself as a quotative 86 times (vs. 286 times for *dizque decir*).

- (16) “*Patrón se enoja conmigo*” *dizque dijo él.*  
 boss 3SG.REFL get.angry with.me said he  
 ‘My boss gets angry with me’, he said.’  
 (Olbertz 2005: 5, ex. 16; spoken corpus cf. Muysken 1978)

A question that arises is whether Ecuadorian *dizque* represents a more advanced stage of grammaticalization by comparison with Bolivian, Colombian and Mexican Spanish. There is some preliminary evidence from Kany (1944) that, in an earlier stage, Ecuadorian Spanish displayed variable scope and a flexible syntactic position in coastal dialects vs. the Highlands.

The Ecuadorian data in Kany (1944) is divided in two sections: coast, Highlands. For the Highlands Kany lists thirteen examples. In each of the thirteen examples *dizque* and its variants immediately precede the verb. This is consistent with the description of Olbertz (2005), which is based on recordings that Peter Muysken did in 1978 in a small town in the province of Cotopaxi (Central Highlands). For the coast, Kany lists nine examples. In five of them *dizque* and its variants immediately precede the verb. In three of them *dizque* is preverbal but not adjacent to the verb (compare with the *dizque* + V reanalysis assumption of 16). The coastal dialect would seem closer to Colombian and Mexican in terms of the syntactic distribution of *dizque*. Determining what *dizque* modifies in each case may be difficult without context. One of these cases is sentence-initial (17) and it seems to modify the clause. Another example modifies a phrase (18).

- (17) *Dizque por arriba todo lo arregl-aban a látigo o a bala.*  
 forup all it fix-3PL.PST by whip or by bullet  
 ‘They say up there they resolved all quarrels with whips or bullets’

<sup>2</sup> Spanish original: “*dizque* + V debe de haberse llegado a reinterpretar como construcción verbal compleja, lo cual explica que la posición de la negación delante del primer elemento de esta construcción”



(Kany 1994: 172, ex. 1; *Don Goyo* by D. Aguilera Malta 1938  
(or. 1933))

- (18) *arriba*, *dizque en el desmonte de un pariente de él*  
up.there in the clearing of a relative of he  
'up there, *supposedly* in a clearing of a relative of his'  
(Kany 1994: 172, ex. 3; *Nuestro Pan* by E. Gil Gilbert)

The same book from which (17) is extracted contains examples in two other positions. Demetrio Aguilera Malta, one of the literary sources for these examples, was born in the coastal city of Guayaquil; so was Enrique Gil Gilbert, the other quoted source. In the book from which (18) is taken there are three additional examples where *dizque* immediately precedes the verb. Both writers demonstrate a flexible syntactic position for *dizque*.

I assume that these differences in syntactic distribution reflect a more advanced stage of grammaticalization in the Highlands. There is a likely cause for the accelerated grammaticalization of *dizque* in the Highlands: Quichua. The coastal examples may reflect an earlier stage. Eight of them seem to modify a sentence, and only one of them a phrase. The most frequent position is immediately before the verb (five examples).

#### 2.4. *Dizque*: A native development or a borrowing?

If *dizque* is undergoing grammaticalization as an evidential and/or epistemic category, a number of related questions arise. First, whether this phenomenon is limited to some dialects or is instead a general characteristic of Spanish. Second, whether it is a recent development or a long-standing process. Third, whether it is a native development or caused/influenced by language contact in Latin America or in Spain. These questions are relevant for the analysis of *dizque* as an emergent evidential or a (pseudo-)grammatical evidential because some Amerindian languages have been described to have grammatical evidentiality, such as Quechua (Faller 2002). The remainder of this section discusses these questions in the order presented.

A comprehensive survey of the geographical distribution of *dizque* in the Spanish-speaking world is still lacking. According to *Diccionario de uso del español de América y España* (Peninsular and Latin American Spanish use dictionary), *dizque* is found in colloquial registers in a vast territory that starts in Mexico, extends to Central America and ends in the Andean region. Recent works include a subset of dialects in this territory: Mexican (Company Company 2004, Olbertz 2007), Panamanian (Lamy 2010), Ecuadorian

(Olbertz 2005), Colombian (Travis 2006) and Bolivian (Babel 2009). Kany (1944) is the closest to a comprehensive geographical survey, but it nears three quarters of a century since its publication. In it *dizque* extends beyond said area to the Caribbean and the Southern Cone. Kany provides examples of *dizque* and its variants for the abovementioned dialects, as well as Argentina, Chile, Peru, Venezuela, Costa Rica, Dominican Republic, Guatemala and the state of New Mexico. Only in four countries the use of *dizque* may have not been reported yet. These are Uruguay, Paraguay, Puerto Rico and Cuba. If *dizque* is not a pan-Hispanic form, it comes close. An anonymous reviewer notes that in the CREA corpus (Modern Spanish Reference Corpus—Spanish *Corpus de Referencia del Español Actual*) *dizque* is primarily attested in Mexico, Dominican Republic and Colombia in written language/fiction. The specialized reader may be familiar with the limitations of CREA as regards Latin American Spanish and oral Spanish.

Regarding its origin, Kany proposes that *dizque* is not an innovation of Latin American dialects—from, say, contact with native languages with grammatical evidentiality such as Quechua, as noted, but rather a continuation from an earlier form in Old Castilian. In view of historical examples and references available to him, Kany concludes that *dizque* was becoming part of the norm by the mid 16<sup>th</sup> c., but it eventually fell out of favor. According to Kany, *dizque* survived in local registers in Spain into the 20<sup>th</sup> c. in areas such as Salamanca and Andalusia. Olbertz (2005, 2007) independently confirms the presence of *dizque* in Old Spanish using CORDE, the online diachronic corpus of the Royal Spanish Language Academy (*Corpus Diacrónico del Español*, Spanish Diachronic Corpus). Thus, *dizque* is potentially a native development, at least in the seed dialect.

If phonologically reduced forms of (1) were commonplace by the 16<sup>th</sup> c., it may be possible that, at that stage, *dizque* featured evidential and epistemic meanings as well. Olbertz (2005, 2007) observes that, in Old Spanish, *dizque* had already acquired evidential and epistemic uses by the late 15<sup>th</sup> century (see also López Izquierdo 2006). Said evidence confirms that *dizque* in Old Spanish had characteristics now attributed to contemporary dialects, strengthening the case for a native development.

Having said that, CORDE has limited representative value across dialects of Spanish. For CORDE is primarily comprised of Spanish authors (that is to say Peninsular, not Latin American). Although we can be fairly certain that *dizque* had an evidential/epistemic use in Spain early on, this does not preclude the possibility of later, parallel developments in Latin American

varieties or borrowing from language contact. Widely spoken languages such as Quechua, which feature grammatical evidentiality (Faller 2002), may have exerted some influence on *dizque* or the forms in (1). For example, contact in the Andean region could have accelerated the evolution of *dizque*, had it been a Peninsular import, or it could have sparked its development. This may have been the case in Bolivian and Ecuadorian Spanish.

More generally, the evolutionary path from a verb of communication to a quotative or reportative evidential is a common cross-linguistic pattern (Aikhenvald 2004: 271). Within the Romance family, we find similar developments for Italian (Giacalone Ramat and Topadze 2007), Galician (Rosales Sequeiros 2001, quoted in Travis 2006), Brazilian Portuguese (Aikhenvald 2007) and French (José Deulofeu, p.c.). This makes the potential contribution of a language, or a language family, more difficult to tease apart from native development.

Yet considering the widespread presence of *dizque* in the Spanish-speaking world, and its apparent roots in the seed dialect, the possible influence of Amerindian languages with grammatical evidential systems is, a priori, relegated to a secondary role, such as favoring or facilitating the change. In this line, Olbertz (2005) reports a higher frequency of use of *dizque* in Ecuador among Quichua dominant bilinguals, over Spanish dominant bilinguals, over Spanish monolinguals<sup>3</sup>. This suggests that the presence of a grammaticalized form to express quotatives and/or hearsay in an Amerindian language may accelerate the grammaticalization of *dizque*.

The thesis of a native development in Castilian still begs the question of whether the seed dialect could have borrowed an evidential from language contact. Basque has been recently described to display some form of grammatical evidentiality (Alcázar 2010, Etxepare to appear). However, the Basque counterpart to *decir*, namely *esan* 'say', does not present advanced signs of grammaticalization (although it is used in reportative and intersubjective contexts). More importantly, there is no research yet demonstrating that Basque had a consolidated reportative and/or quotative prior to the 16<sup>th</sup> c. For this reason, the hypothesis that Castilian could have borrowed an evidential from Basque, or be influenced by Basque evidentiality, while a theoretical possibility, lacks sufficient empirical evidence at this point; and it is otherwise unnecessary in light of parallel developments in Romance.

---

<sup>3</sup> For a recent overview of Spanish in contact with Amerindian languages, including evidentiality, see Klee and Lynch (2009, Ch. 4 and references therein).

In sum, Peninsular Spanish shows an evidential/epistemic use of *dizque* by the 15<sup>th</sup> c. This form may have crossed over the Atlantic, or independent evolutions may have taken place, with or without language contact.

### 2.5. Summary

*Dizque* is a phonological reduction of the prototypical verb of saying in the third person present indicative and the short passive. *Dizque* seems to be in a process of grammaticalization towards an evidential and/or epistemic category. The semantics and syntactic distribution of *dizque* across the Spanish dialects considered in this paper is diverse. The origin of *dizque* is likely a native development of the seed dialect; Peninsular Spanish, where it is now virtually obsolete. That said, *dizque* could have arisen independently in Latin American dialects of Spanish through parallel development for similar evolutions are attested in Romance and elsewhere. It is likely that contact with Amerindian languages that feature evidentiality can accelerate the grammaticalization of *dizque*, particularly in bilingual populations.

## 3. ANALYSIS

*Dizque* has been classified as an emergent evidential (Aikhenvald 2007) on account that it lacks primary evidential meaning and it is optional. More generally, *dizque* is not part of verbal inflection or a verbal clitic (Ecuadorian Spanish being excepted cf. ex. 16)—vs. a majority of languages with grammatical evidentiality (Aikhenvald 2004). Syntactically, *dizque* behaves more like a particle in that it does not seem to observe syntactic restrictions. For instance, in Mexican Spanish, according to Olbertz (2005), *dizque* can modify clauses, phrases and heads; usually, by being adjacent to what it modifies. This last characteristic would suggest that *dizque* is not the exponent of a clausal grammatical category, such as tense, modality or evidentiality. Aikhenvald notes, nonetheless, that Mexican *dizque* meets the four parameters provided by Giacalone Ramat and Topadze (2007) to signal grammaticalization. Compared to *decir* ‘to say’, *dizque* has gained positional freedom and variable scope as it has undergone decategorization and semantic bleaching. But it best analyzed as an emergent evidential for Aikhenvald in light of the characteristics mentioned at the beginning.

As we inspect different dialects of Spanish, however, some of the characteristics of *dizque* as an emergent evidential are unexplained. In the case of Bolivian Spanish, for example, epistemic interpretations seem to be relegated to the pragmatic domain. In Mexican and Colombian, when *dizque* modifies sentences, the interpretation appears to be evidential at the semantic level. Some of these findings are not unique to *dizque*. Regarding emergent evidentials more generally, evolutions parallel to *dizque* are also attested in Galician, Romanian, Sardinian, and Sicilian (Cruschina and Remberger 2008). In these languages, the evidential meaning appears to be semantic and epistemic meaning pragmatic, and seemingly not too salient. Cruschina and Remberger (2008) compare the semantics of Colombian *dizque* with its European counterparts: ‘the labeling function [the epistemic interpretation reported by Travis (2006)] is less easy to find - perhaps even impossible - in the varieties under discussion. It is also not yet clear whether each variety has a genuine dubitative [...] or if the dubitative interpretation is due to pragmatic implicatures.’ (p. 13).

Some scholars note increases in the frequency of explicit marking of evidentiality with *dizque* (Babel 2009 for Bolivian), which suggest that evidentiality is becoming an obligatory grammatical category. Emergent evidentials are not obligatory in root declarative sentences and thus fail to meet one of the criteria for the narrow definition of evidentiality. However, not all languages believed to have grammatical evidentiality meet this requirement. For example, evidential markers in Quechua are not obligatory (Faller 2002).

Other characteristics of *dizque* are also reminiscent of grammatical evidentiality. For instance, there are collocations of *dizque* with the verb *decir* in quotative uses (Olbertz 2005, Babel 2009), which are also found in languages with grammatical evidentiality (Aikhenvald 2004). In Ecuadorian Spanish, *dizque* has acquired a fixed syntactic position before the verb. Furthermore, it appears that in this dialect *dizque* can no longer modify phrases or heads. Sentential scope is the expected scope of a sentential grammatical category.

This presents a problem for the characterization of *dizque* as an emergent evidential in Bolivian and Ecuadorian Spanish, as well as for equivalent forms in some Romance languages. In particular, the more epistemic, syntactically flexible versions of *dizque* attested in Mexican and Colombian Spanish need to be explained away. The related question arises of whether the tendencies between clausal modification with evidential

interpretation and phrasal/head modification with epistemic interpretations are anecdotal, or language particular, vs. representative of a likely path of grammaticalization. While we can integrate the flexible syntactic position of *dizque* into a process of grammaticalization, the coexistence of evidential and epistemic interpretations, and relative dominance predicated on their scope, may have not been reported before. This does not mean that the data is suspect for, according to both Squartini (2007b) and Aikhenvald (2007), there is little knowledge about the intermediate phases of grammaticalization of evidentiality.

I propose that emergent evidentials could go through an *epistemic phase* as part of their evolution from an emergent evidential to a grammatical evidential due to subjectification. Following Traugott's adoption of Lyons' definition, subjectivity is 'the locutionary agent's expression of himself and his own attitudes and beliefs' (Lyons 1982: 102 via Traugott 2010: 33). Subjectification is for Traugott the mechanism by which "meanings are recruited by the speaker to encode and regulate attitudes and beliefs" (2010: 35). Although subjectification is independent of grammaticalization, both are strongly correlated in primary grammaticalization (Traugott 2010: 40). Under my proposal, epistemic interpretations of emergent evidentials would not count against their consideration as an evidential expression in the realm of grammatical evidentiality.

The syntactic category of *dizque* does not seem to be unitary given the contrast between its syntactic ubiquity in some dialects and its fixed position in others. The general tendency seems to be for *dizque* to gravitate to the immediate preverbal position. To resolve said quandary in the analysis of *dizque*, I propose that there are two different syntactic categories for *dizque*. *Dizque-1* would be a particle with variable placement in the clause and able to modify a diversity of constituents. *Dizque-2* would be the exponent of a clause-level grammatical category (Cinque 1999, Speas and Tenny 2003). As such, *dizque-2* would be recognizable in its scope being limited to clauses (vs. phrases and heads); a (pseudo-)fixed syntactic position; being obligatory if reanalyzed as part of the (finite) verb of the main clause; and in exhibiting collocations (e.g., with *decir* 'say' in quotative uses).

*Dizque-1* and *dizque-2* can seemingly coexist in the same dialect in what appears to be an overlapping relation (Hopper 1991). If they did, this could serve as an explanation for the tendencies reported by Travis and Olbertz for Colombian and Mexican Spanish, respectively, which suggest two different semantic domains of evaluation for seemingly the same linguistic expression.

However, if Ecuadorian Spanish would signal the eventual behavior of *dizque* as an evidential category, it is noteworthy that this seemingly overlapping relation appears to end with the loss of *dizque-1* in favor of *dizque-2*. The expression of evidentiality in Basque dialects spoken north/south of the French-Spanish border (Alcázar 2010, Etxepare to appear) also suggests that particle-like emerging evidentials evolve into a verbal clitic with a fixed syntactic position and clausal scope.

#### 4. CONCLUSION

This paper has focused on the grammaticalization of Latin American *dizque*, one of the few emergent evidentials on which recent papers have been published. Emergent evidentials seem to be on the frontline of whether a continuum exists between grammatical and lexical evidentiality. A broader look at the semantics and syntactic distribution of *dizque* across dialects of Latin American Spanish reveals a more complex picture than previously acknowledged. In some dialects, *dizque* appears to be closer to the behavior of an optional grammatical evidential. I have proposed that *dizque* may be analyzed as two syntactically different linguistic expressions: one a particle; the other the exponent of a clause-level projection. Future research on emergent evidentials needs to establish the tendencies reported for Latin American *dizque* are either language-specific, or representative of a path of grammaticalization found in other languages.



## REFERENCES

- Aikhenvald, Alexandra Y. (2007): "Information source and evidentiality: what can we conclude?", in Mario Squartini (ed.) *Evidentiality between lexicon and grammar*. *Rivista di Linguistica* 19 (2), pp. 209-227.
- (2004): *Evidentiality*. Oxford: Oxford University Press.
- Alcázar, Asier (2010): "Information source in Spanish and Basque: A parallel corpus study", in Gabriele Diewald and Elena Smirnova (eds.): *Linguistic realization of evidentiality in European languages*. Berlin, New York: Mouton De Gruyter, pp. 131-156.
- Anderson, Lloyd B. (1986): "Evidentials, paths of change, and mental maps: typologically regular asymmetries", in cpp. 273-312.
- Biber, Douglas and Edward Finegan (1989): "Styles of Stance in English: Lexical and Grammatical Marking of Evidentiality and Affect", *Text* 9, pp. 93-124.  
<http://dx.doi.org/10.1515/text.1.1989.9.1.93>
- Brown, Penelope and Stephen C. Levinson (1987): *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Bybee, Joan (2003): "Mechanisms of change in grammaticization: The role of frequency", in Brian D. Joseph and Richard. Janda (eds.): *The Handbook of Historical Linguistics*. Oxford: Blackwell, pp. 602-623.
- Chafe, Wallace L. (1986): "Evidentiality in English Conversation and academic writing", in Wallace Chafe and Johanna Nichols (eds.): *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex, pp. 261-72.
- and Johanna Nichols (eds.) (1986): *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.



- Cinque, Guglielmo (1999): *Adverbs and functional heads: a cross-linguistic perspective*. Oxford: Oxford University Press.
- Cruschina, Silvio and Eva Remberger (2008): "Hearsay and reported speech: Evidentiality in Romance", in Paola Benincà, Federico Damonte and Nicoletta Penello (eds.): *Selected Proceedings of the XXIV Incontro di Grammatica Generativa, Rivista di Grammatica Generativa* 33, pp. 95-116.
- Diewald, Gabriele and Elena Smirnova (eds.) (2010): *Linguistic Realization of Evidentiality in European Languages*. Berlin, New York: Mouton de Gruyter.  
<http://dx.doi.org/10.1515/9783110223972>
- Etxepare, Ricardo (to appear): "Omen bariazioan", *Euskaren sintaxia, Anuario del Seminario de Filología Vasca 'Julio de Urquijo'/ International Journal of Basque Language and Linguistics*.
- Faller, Martina (2002): *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. PhD Dissertation. Stanford University.
- Friedman, Victor (1986): "Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian", in Wallace Chafe and Johanna Nichols (eds.): *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex, pp. 168-187.
- Giacalone Ramat, Anna and Manana Topadze (2007): "The coding of evidentiality: a comparative look at Georgian and Italian", in Mario Squartini (ed.) *Evidentiality between lexicon and grammar. Rivista di Linguistica* 19 (2), pp. 7-38.
- Givón, T. (1982): "Evidentiality and Epistemic Space", *Studies in Language* 6 (1), pp. 23-49.  
<http://dx.doi.org/10.1075/sl.6.1.03giv>
- Hopper, Paul J. (1991): "On some principles of grammaticalization", in Elizabeth Closs Traugott and Bernd Heine (eds.): *Approaches to Grammaticalization*. Volume I. Amsterdam: John Benjamins, pp. 17-35.  
<http://dx.doi.org/10.1075/tsl.19.1.04hop>
- Jacobsen, William H. (1986): "The Heterogeneity of Evidentials in Makah", in Wallace Chafe and Johanna Nichols (eds.): *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex, pp. 3-28.
- Lamy, Delano (2010): "Marcador epistémico y discursivo: 'dizque' en el español panameño." Paper delivered at WSS5, University of North Carolina.

- Lazard, Gilbert (2001): "On the grammaticalization of evidentiality." *Journal of Pragmatics* 33, pp. 358-68.  
[http://dx.doi.org/10.1016/S0378-2166\(00\)00008-4](http://dx.doi.org/10.1016/S0378-2166(00)00008-4)
- López Izquierdo, Marta (2006): "L'émergence de dizque comme stratégie médiative en espagnol médiéval." *Cahiers d'études hispaniques médiévales* 29, pp. 483-493.
- Mithun, Marianne (1986): "Evidential Diachrony in Northern Iroquoian", in Wallace Chafe and Johanna Nichols (eds.): *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex, pp. 89-112.
- Nuyts, Jan (2001): *Epistemic modality, language, and conceptualization: A cognitive-pragmatic perspective*. Amsterdam: John Benjamins.  
<http://dx.doi.org/10.1075/hcp.5>
- Olbertz, Hella (2007): "Dizque in Mexican Spanish: the subjectification of reportative meaning", in Mario Squartini (ed.): *Evidentiality between lexicon and grammar*. *Rivista di Linguistica* 19 (2), pp. 151-172.
- (2005): "Dizque en el español andino ecuatoriano: conservador e innovador", in Hella Olbertz and Pieter Muysken (eds.): *Encuentros y conflictos. Bilingüismo y contacto de lenguas en el mundo andino*. Madrid: Iberoamericana, pp. 77-94.
- Oswalt, Robert L. (1986): "The evidential system of Kashaya", in Chafe and Nichols: pp. 29-45.
- Rosales Sequeiros, Xosé (2001): "Historical evolution of attributive adverbs in Galician". Paper presented at the 15<sup>th</sup> International Conference on Historical Linguistics, Melbourne.
- Speas, Margaret and Carol Tenny (2003): "Configurational properties of point of view roles", in Anna Maria Di Sciullo (ed.): *Asymmetry in Grammar. Volume I: Syntax and Semantics*. Amsterdam/Philadelphia: John Benjamins, pp. 315-344.
- Squartini, Mario (ed.) (2007a). *Evidentiality between lexicon and grammar*. *Rivista di Linguistica* 19 (2).
- (2007b) "Investigating a grammatical category and its lexical correlates", in Mario Squartini (ed.) *Evidentiality between lexicon and grammar*. *Rivista di Linguistica* 19 (2), pp. 1-6.

- Traugott, Elizabeth Closs (2010): "(Inter)subjectivity and (inter)subjectification: a reassessment", in Kristin Davidse, Lieven Vandelanotte, Hubert Cuyckens (eds.): *Subjectification, Intersubjectification and Grammaticalization*. Berlin, New York: Mouton de Gruyter, pp. 29–69.  
<http://dx.doi.org/10.1515/9783110226102.1.29>
- and Richard Dasher (2002): *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
- Travis, Catherine (2006): "Dizque: a Colombian evidentiality strategy", *Linguistics* 44, pp. 1269-1297.  
<http://dx.doi.org/10.1515/LING.2006.041>
- Willet, Thomas (1988): "A cross-linguistic survey of the grammaticization of evidentiality", *Studies in Language* 12, pp. 51-97.  
<http://dx.doi.org/10.1075/sl.12.1.04wil>



## **Decir en español rioplatense: un análisis semanto-pragmático**

| Alicia M. Ocampo  
| *University of Minnesota*

**Abstract:** This article deals with the expression of evidentiality in Rioplatense Spanish, a variety spoken in the area of Buenos Aires, Argentina. The analysis focuses on *decir* ('to say'), a verb used to express most evidential categories. Special attention is given to the form *dice*, which is in the process of semantic and grammatical changes. *Dice* moved from lexicon and grammar into discourse to become a particle. In its most advanced stage of transformation it is reinterpreted as a hearsay index, with encoded evidential meaning.

**Keywords:** Evidentiality, evidentials, hearsay, verb *decir*, discoursivization, grammaticalization, semantic change, Rioplatense Spanish.



## 1. INTRODUCCIÓN

El presente trabajo se centra en el análisis de los diversos usos evidenciales del verbo *decir* que se registran en el español coloquial del Río de la Plata. Este estudio se ha basado en datos correspondientes a un corpus obtenido en La Plata, Argentina, entre los años 1986 y 1997. Se trata de conversaciones informales (charlas de sobremesa, conversaciones de café) grabadas por la investigadora a 20 hablantes nacidos en La Plata (amigos y familiares). Las transcripciones utilizadas corresponden a un total de 20 horas de grabación.

*Decir* es usado principalmente para expresar evidencialidad. En su concepto más amplio, esta noción se relaciona con la modalidad epistémica, y se refiere al modo en que los hablantes marcan la fuente de la información que están comunicando. Determinada información puede presentarse como verdadera porque se ha presenciado el hecho reportado, o porque una autoridad lo afirma, o porque se tiene evidencia directa del asunto. O bien, la información puede presentarse como posible o probable, porque se conoce por vía indirecta, o porque sólo se supone o se infiere.

En este trabajo se intenta demostrar una hipótesis en relación a la forma *dice*: los datos del rioplatense indican que se encuentra en un proceso de transformación de verbo a partícula evidencial *de oídas*. Según nuestra interpretación, esto representa un movimiento hacia la discursivización.

Primero es necesario presentar una categorización de la evidencialidad, con ejemplos que aclaren los distintos conceptos. En segundo lugar, se listarán las distintas formas de *decir* que los hablantes usan para expresar las subcategorías evidenciales. Más adelante se analizan ejemplos indicadores de que la forma *dice* se está moviendo de verbo a partícula discursiva evidencial. Por último, se da el marco teórico para interpretar el cambio observado en los datos (presentados en la sección anterior) como un movimiento hacia la discursivización.

## 2. CATEGORIZACIÓN DE LA EVIDENCIALIDAD

Aikhenvald (2004) sugiere que, a pesar del reciente interés en los estudios de evidencialidad, todavía sigue siendo una de las categorías gramaticales menos conocidas. Según reporta Aikhenvald (2006), el término 'evidencial' fue introducido por primera vez por Jakobson (1957). Sin

embargo, aparentemente, ya había sido Boas (1938) el primero que hablara de lenguas que marcan obligatoriamente la fuente de la información como adquirida en forma visual, en forma auditiva o por inferencia.

Papafragou et al. (2007) han investigado la adquisición de la evidencialidad y su relación con el razonamiento de los niños en cuanto a conceptos evidenciales. Aparentemente, se puede asumir que las habilidades que los humanos tienen para controlar la fuente de la información surgen en forma independiente, y anteceden a los distintos aspectos de los sistemas lingüísticos que codifican dicha fuente. Según estos autores, en los seres humanos existe una predisposición cognoscitiva a controlar el origen de la información que reciben, y a comparar y evaluar las fuentes. Por último, las diferencias que pudieran existir en cuanto a la codificación lingüística de la evidencialidad no afectan substancialmente el razonamiento sobre el acceso de la información que los niños pequeños pudieran tener.

Considerando distintos estudios de evidencialidad como categoría gramatical en una variedad de lenguas, Aikhenvald (2004) afirma que la evidencialidad, la modalidad y el modo corresponden a categorías completamente distintas. A veces, existen extensiones semánticas de ciertas categorías evidenciales que presentan superposición con algún tipo de modalidad (probabilidad, posibilidad, o la evaluación del hablante en cuanto a la confiabilidad de la información, por ejemplo), lo que depende de cada sistema en que ocurre. Pero, según Aikhenvald, esto no convierte a la evidencialidad en una subcategoría de alguna modalidad, sino que puede considerarse una categoría independiente. Por otra parte, después de estudiar las coincidencias y diferencias entre marcadores de evidencialidad y de modalidad epistémica, Speas (2008) concluye que existen diferencias sistemáticas entre ellos. Sin embargo, no queda claro si esto significa que son completamente distintos, o si los evidenciales constituyen un tipo especial dentro del modo epistémico.

Por su lado, Chafe (1986) estudia las marcas evidenciales en inglés, y halla diferencias en el uso de éstas entre la lengua coloquial y la lengua escrita con propósitos académicos. En su estudio taxonómico de los fenómenos evidenciales, no usa el término *evidencialidad* en un sentido restringido a la expresión exclusiva de la ‘evidencia’; lo aplica, en cambio, en un sentido amplio que cubre todas las actitudes que un hablante tiene en relación al conocimiento de la información.

Givón (1989) redefine las modalidades epistémicas tradicionales en términos discursivo-pragmáticos. Así distingue entre presuposiciones,

aseveraciones ‘realis’, aseveraciones ‘irrealis’ y aseveraciones negativas. Son principalmente las aseveraciones ‘realis’ las relevantes para los usos de *decir*, porque son éstas las que más tienen que ver con la evidencialidad. En ellas, la información es fuertemente aseverada, y el hablante debe indicar de qué fuente de evidencia dispone para sostenerla. De acuerdo a Givón (1989), las fuentes más típicas serían: (a) la experiencia directa, (b) la experiencia de otros que ha sido comunicada, y (c) la inferencia (producto de la experiencia directa o comunicada).

Existen muchas lenguas en las que la evidencialidad se presenta codificada en morfemas específicos. Como categoría gramatical, Aikhenvald (2004) estima que estos evidenciales existen aproximadamente en una cuarta parte de las lenguas del mundo. En áreas de Norte y Sudamérica existen lenguas que tienen esta característica: el wintu (Schlichter 1986), el maricopa (Gordon 1986), el quechua (Weber 1986) y las lenguas jaqi (Hardman 1986), entre otras. También en la región de los Balcanes (Friedman 1986), y en lenguas asiáticas como el japonés (Aoki 1986), el tibetano (DeLancey 1986) y el turco (Aksu-Koç y Slobin 1986).

El español es diferente a estas lenguas, ya que no indica las distintas categorías evidenciales por medio de marcas gramaticales. En la mayoría de los casos, en español el hablante puede optar por indicar la evidencialidad de un modo explícito, y para ello dispone de distintos instrumentos lingüísticos que se usan como marcas evidenciales: adverbios, elementos léxicos y verbos utilizados para discurso indirecto, entre otros. Bermúdez (2005) considera que hay lenguas que solamente registran *estrategias* de evidencialidad: las que la indican por medio de partículas que están fuera de un sistema, o por elementos no morfológicos, o por elementos gramaticales opcionales. En este sentido, el español sería una de ellas.

El presente estudio toma la evidencialidad en su interpretación amplia, del mismo modo que Chafe (1986). Por ello, se ha encontrado apropiado basar el análisis, en su mayor parte, en el cuadro que Chafe (1986) propone, que incluye distintas subcategorías comprendidas en el concepto de *evidencialidad*, y cómo éstas se relacionan entre sí.

Con excepción de las experiencias directas, distintos usos de *decir* cubren un espectro muy amplio del resto de funciones necesarias para expresar los distintos modos del conocimiento y otros tipos de información evidencial. En los datos se ha encontrado el caso interesante de la forma invariable *dice*, que llega a utilizarse para cualquier persona gramatical, y con distintos valores temporales y aspectuales.



El Cuadro 1 representa una taxonomía de la evidencialidad que, aunque se basa en el de Chafe (1986), incluye cambios inspirados en parte en la clasificación de Givón (1989) (a la que se hizo referencia más arriba), y otros propuestos por la autora.

EVIDENCIALIDAD

fuentes del conocimiento →	modos del conocimiento	la información en contraste con
		[no confiable]
		↓
		↓
		↓
hablante/ fuente impersonal →	creencias, opiniones →	C
		O ↓
		N
		O →
experiencia/ comunicación →	inferencias	C ↓
		I
		M
lenguaje →	de oídas →	I ↓
		E →
evidencia sensorial →	experiencias directas →	N ↓
		T
		O ↓
		[confiable]

Cuadro 1. Categorización de la evidencialidad

En el centro del cuadro aparece la palabra *conocimiento*, escrita en forma vertical. De acuerdo a Chafe, es la información básica cuyo estatus es clasificado de un modo u otro por los marcadores de evidencialidad. Al *conocimiento* o *información* puede aplicarse una *escala de confiabilidad* que va de *no confiable* a *confiable*. En el cuadro también se representan distintos *modos del conocimiento* que tienen origen en fuentes diversas. Hay una gradación de veracidad entre las *creencias* (menos veraces) y las *experiencias directas* (más veraces), pero esta gradación es independiente de la *escala de confiabilidad* aplicada al *conocimiento*. Por otra parte, los distintos *modos del conocimiento* tienen fuentes identificables. Las *creencias* y *opiniones* se originan en el mismo hablante o en una fuente impersonal; las *inferencias* pueden fundarse en la experiencia misma o en la comunicación; el conocimiento obtenido *de oídas* tiene por fuente el lenguaje; y las *experiencias directas* tienen origen en la evidencia sensorial de los hechos.

Para Chafe (1986), los distintos *modos del conocimiento* implican diferencias en grado de *confiabilidad*. Por ejemplo, las *opiniones* y las

*creencias* aparecen en la escala como *menos confiables*, en oposición a las *experiencias directas*, mucho *más confiables* por ser obtenidas gracias a los sentidos. Sin embargo, la gradación en cuanto a *confiabilidad* no implica necesariamente distintos *modos del conocimiento*, sino que puede darse dentro de un mismo modo.

En relación a la columna derecha del cuadro, según Chafe el hablante puede contrastar el conocimiento de la *información presentada* con los *recursos verbales* o las *categorías semánticas* que él mismo elige para comunicar esa información. Si el hablante no encuentra la palabra exacta para expresar lo que desea, puede indicar ese desajuste usando alguna marca evidencial. Por último, es posible indicar contraste entre la *información comunicada* y la *información esperada*, según las *expectativas* originadas en la norma cultural del hablante, o según *expectativas* producidas por el contexto mismo en el que se presenta la información.

A continuación, se intentará clarificar más el Cuadro 1 mediante la presentación de algunos ejemplos. En primer lugar, los distintos *modos del conocimiento*:

### **Creencias/opiniones**

(a) *Para mí* que Juan ya no viene.

La información ‘Juan ya no viene’ está marcada como una opinión personal del hablante. En este caso, *para mí* indica que se trata de una creencia poco confiable.

### **Inferencia**

(b) Mañana *debe ser* el último día para entregar el formulario.

Para inferir la aludida información, el hablante se basa en datos no revelados, que puede haber obtenido por su propia experiencia o por algo que le dijeron. *Debe ser* es aquí la marca evidencial de inferencia, no muy confiable.

### **Conocimiento de oídas**

(c) Me *dijo* que mi composición era muy buena.

El hablante podría ser un estudiante que no está muy seguro de la calidad de su trabajo. La información de que su composición es muy buena la obtiene por medio de una comunicación con su profesor. “Me *dijo*” indica: conocimiento *de oídas*, bastante confiable, originado en el lenguaje.

**Experiencia directa**

(d) *Veo* que te has cortado el pelo.

Aquí, un tipo de evidencia sensorial, la vista, provee el conocimiento sobre el corte de pelo de la otra persona. El verbo *ver* funciona como indicador evidencial de experiencia directa, o sea, muy confiable.

En cuanto a la sección del Cuadro 1 reservada a *contraste de la información*, se verán los siguientes ejemplos:

**Contraste entre el conocimiento de la información dada y los recursos verbales utilizados:**

(e) Y entonces, en ese momento, aparece *algo así como* un monstruo cibernético.

El hablante no está muy convencido de que su categorización sea la más adecuada para lo que intenta describir. Por eso, por medio de *algo así como*, indica que los recursos verbales a los que recurre son sólo aproximados.

**Contraste entre el conocimiento de la información dada y las expectativas:**

(f) *Es increíble* la desvergüenza de esos comerciantes. Son verdaderos piratas.

Al utilizar *es increíble*, el hablante está marcando que el comportamiento de los comerciantes referidos no es el esperado por la sociedad, sino que se desvía de la norma.

Finalmente, la *escala de confiabilidad* que pueda aplicarse a la información comunicada se ilustra en los siguientes tres ejemplos, pertenecientes todos al *modo del conocimiento* de tipo *creencia/opinión*:

(g) *No parece* que le vayan a dar el premio.

(h) *Quizás* le den un premio.

(i) *Es indudable* que la obra merece un premio.

Estas expresiones difieren en el grado de confianza que el hablante tiene en cuanto a la información que comunica. En (g) se presentan datos no confiables, mientras que en (i) la información está percibida casi como un hecho. Las tres marcas evidenciales, dadas por: *no parece* + Subjuntivo en (g); *quizás* + Subjuntivo en (h); y *es indudable* + Indicativo en (i), producen la gradación.

### 3. DECIR: VERBO EVIDENCIAL

Dentro de la clasificación presentada, en rioplatense *decir* puede ser usado para indicar la mayor parte de las subcategorías que se incluyen en la noción amplia de evidencialidad. Entre ellas, como podría esperarse, la única que no puede ser expresada por *decir* es la referida a experiencias directas, obtenidas a través de los cinco sentidos.

Así, la Tabla 1 anticipa los usos de *decir* encontrados en los datos del rioplatense para cada una de las subcategorías del Cuadro 1. Estos usos se encontrarán ejemplificados en el apartado siguiente.

- I. modos del conocimiento:
  - I.1. creencias y opiniones
    - I.1.1. fuente de conocimiento: el hablante: ***digo yo***
    - I.1.2. fuente de conocimiento: impersonal: ***como dicen***
  - I.2. inferencias; fuentes: por experiencia o comunicación: ***digo yo***
  - I.3. conocimiento por la lectura o *de oídas*; fuente: el lenguaje
    - I.3.1. fuente de conocimiento: lenguaje escrito: ***dice***
    - I.3.2. fuente de conocimiento: lenguaje oral: ***dice, dijo, digo, dije***, etc
- II. escala de confiabilidad
  - II.1. información poco confiable: ***digamos, digo yo***
  - II.2. información muy confiable: ***te digo***
- III. contraste entre la información dada y los recursos verbales, o las expectativas.
  - III.1. contraste con recursos verbales: ***te diría, no te diría, no te digo***
  - III.2. contraste con expectativas: ***te diré, te digo***

**Tabla 1.** Usos de decir encontrados en los datos del rioplatense

Es importante notar que algunas de estas expresiones se utilizan para comunicar más de un tipo de evidencialidad. En el caso de *digo yo*, dado que se usa para marcar la información como *poco confiable* (dentro de la *escala de confiabilidad*), resulta lógico usarla también tanto para las *creencias/opiniones* originadas en el mismo hablante, como también para las *inferencias*, ya que ambas pueden considerarse *poco confiables*.

En el caso de *te digo*, los hablantes quieren marcar un *contraste con expectativas*, y al mismo tiempo indicar que, aunque la información provista no es la esperada, los oyentes deben creerla porque es *muy confiable*.

Veremos a continuación una serie de ejemplos encontrados en los datos que ilustran los usos de *decir* incluidos en la Tabla 1. Los números refieren a las categorías y subcategorías dentro de esta tabla.

## I. MODOS DEL CONOCIMIENTO

Los ejemplos (1-2) ilustran I.1. creencias y opiniones del mismo hablante, o de una fuente impersonal:

### I.1. Creencias y opiniones

#### I.1.1. Fuente de conocimiento: el hablante

En (1) la hablante P relata una anécdota de su vecina, casada, que había estado en una lucha física con un amigo en ausencia de su marido. Por sus palabras, resulta obvio que P encuentra esto reprochable y expresa su opinión (*digo yo*) al preguntarse qué pasaría si el marido hubiera llegado en ese momento:

- (1) P: dijo don Víctor ¿está bien señora? o algo así porque no sé qué había tenido ella - si yo es - si estoy fuerte dice, si: () ho si hoy dice estuve luchando ¿cómo se llama? ah Felipe () Felipe es el () esposo de la señora de Santillán que vive ahí por la [calle] 505. Dice, estuve - estuve luchando con Felipe y no pudo él conmigo () y nosotros si no dice don Víctor () estuvo () luchando con Felipe y si en ese momento entra el esposo y la ve luchando con Felipe y () y **digo yo**, qué sé yo qué - qué pasará ¿no ves?

A: era un escándalo un escándalo

P: no no no no asumía la () su posición de: () de de () de mujer de mujer que () un poco

L: decente

P: qué sé yo

En (1), la fuente es la misma hablante, que invoca como *modo de conocimiento* su propia opinión: *digo yo* equivale a *en mi opinión*. Sirve para presentar lo que piensa un hablante, e indica además mediana confiabilidad en lo que expresa.

### I.1.2. Fuente de conocimiento: impersonal

En el ejemplo siguiente no se individualiza la fuente. El hablante F no está muy convencido de lo que afirma, por eso le interesa marcar que él en realidad está haciéndose eco de una creencia generalizada (fuente impersonal):

(2) F: estamos en democracia **como dicen**

El ejemplo (3) presenta el caso de una inferencia:

### I.2. Inferencias

La expresión *digo yo* indica una inferencia, derivada de alguna situación experimentada por el hablante, o de información obtenida a través de la comunicación. En (3) se trata de una inferencia originada por experiencia directa.

El hablante C es bioquímico y narra un incidente que le ocurrió una noche que le tocaba hacer guardia en un hospital. En el relato, el hablante asume (erróneamente) que si un policía lo busca ha de ser porque le han robado o chocado el auto. La confiabilidad en lo que infiere es baja:

(3) C: e:s bravo, es bravo. Yo te digo, a mí me han llamado, suponete, este último: - lo que pasa es que: vos ves en estas - cosas en otros lugares, que no ocurren porque está todo más ordenado. Me llamaron, en esta última guardia no, la anterior creo que fue, me llaman a la:s doce y media de la noche, y me dicen va a venir un señor de Florencio Varela<sup>1</sup>, un policía de Florencio Varela, (0.4) este::, con una muestra para que le den una alcolemia, - alcolem - alcohol en la sangre, en un tipo que - estuvo en un accidente de tránsito. (1) Yo digo no, eso aquí no se hace. Vaya a un hospital que t (0.8) que: que tiene guardias para atender a agu:dos. Eso es una cosa: que podrían derivarlo a un laboratorio de Policía.

F: mhm

C: bueno, entonces me dice:n, bueno sí, pero el policía va a venir. Hablo yo y le digo no:, no venga porque acá no se hace. (0.6) Entonces, bueno, pasó. Yo pensé que se había superado. Como a las tres y media de la mañana, me vuelven a llamar al laboratorio y me dicen: (0.8) este, aquí

<sup>1</sup> Florencio Varela es una localidad de la Provincia de Buenos Aires.

en terapia intensiva hay un:: policía que lo busca. Bue  **digo yo me afanaron<sup>2</sup> el a::uto:, chocaron con mi a::uto y m: m m me veo en un despiole<sup>3</sup>**. Voy a hablar con el tipo y me dice, ah, yo vengo a buscar la mue:stra de sangre – ¡ino!!

Los ejemplos (4-5) ilustran situaciones del tipo **I.3.** el conocimiento ha sido obtenido por medio del lenguaje:

### **I.3. Conocimiento por la lectura o de oídas; fuente: el lenguaje**

#### I.3.1. Fuente de conocimiento: lenguaje escrito

En (4) el hablante se refiere a la etiqueta de un frasco de vitaminas que provee las indicaciones para el paciente:

(4) M: la vitamina ésa que es que la ¿qué  **dice**  acá? ¿'no masticar'  **dice**  acá donde  **dice** -?

El verbo *dice* (o sea, *la etiqueta dice*) es un uso metafórico atribuido a una entidad que no puede hablar o producir sonidos lingüísticos. Debe interpretarse como 'se lee'.

#### I.3.2. Fuente de conocimiento: lenguaje oral

En (5), M explica que, según un médico, las vitaminas son más beneficiosas si se las toma en diferentes momentos a través del día.

(5) M:no: bueno pero es - pero es un docto:r que enseña de los dolores reumá:ticos (y)  **dice**  que las vitaminas que son necesarias. [cita las palabras del doctor] Lo que le conviene, cuando se toma vitamina, este: - repártalas en varias tomas. En vez de tomarse una toda una vez (0.8) la reparte en varias tomas

A: a mí me parece que es más efectivo. Y siempre  **dice**  tomar las vitaminas después de comer. Así se fijan más.

Éste es un ejemplo prototípico del uso de *decir* en discurso indirecto. Dentro del tipo **I.3.2.** es interesante la forma *dice*, que parece estar en proceso de discursivización con función de marcador evidencial *de oídas*. Esto se verá en detalle más adelante.

<sup>2</sup> En rioplatense, 'afanar' significa 'robar'; 'despiole' se interpreta como 'desastre'.

<sup>3</sup> Lo marcado en bastardilla indica la inferencia del hablante.

## II. ESCALA DE CONFIABILIDAD

En cuanto a **II.**, a continuación (6-8) ejemplifican casos de baja y alta confiabilidad en la información:

### II.1. Información poco confiable

En (6), la hablante L no encuentra la palabra exacta que quiere expresar, entonces propone una definición para que los otros participantes la ayuden a encontrar tal palabra. Sin embargo no confía mucho en que su definición sea la más apropiada. Esto lo indica mediante el uso de *digamos*:

(6) L: leía o::h leía ta:nto () y: bueno, todos lo: - él casi era como una **digamos** ¿cómo se llama ese lugar donde van todos los los este: () los y **digamos** los inmigrantes que se van a: a asesorar de cosas? allá en casa era: () una ¿cómo te diría? u - una embajada no, ya es muy - muy grande muy grande decir [...]

M: una embajada

L: ¿una qué?

A: [...] un consulado

F: un consulado

L: un consulado una cosa así, porque todos los p:apeles se los hacía papá.

De este modo, otro hablante propone la palabra *consulado*, que parece ser más aproximada para el contexto dado.

En (7) vemos el uso de *digo yo*, que también marca bajo grado de confiabilidad en lo que el mismo hablante dice:

(7) P: sí, es medio terrorífico el tipo, eh. Te digo que es medio: [loco]

S: un día estaba haciendo gimnasia. Había tirado la cerca, y haciendo gimnasia como loco con pantaloncitos cortos ahí. Como loco **digo yo**

A: he he he .

F: como lo que era



En este ejemplo P y S relatan sobre unos linyeras<sup>4</sup> con deficiencias mentales que habían ocupado una casa abandonada de la vecindad. Uno de ellos parecía fanático por la gimnasia y en pleno invierno hacía ejercicios en shorts, en un despliegue de energía. En (7) hay un juego de palabras, como se verá. La expresión ‘hacer gimnasia *como loco*’ significa ‘frenéticamente’. S evalúa el modo de actuar del linyera, implicando que se comportaba ‘como loco’. Sin embargo, de inmediato usa la expresión *digo yo* para hacerles saber a los oyentes que ésta es una opinión suya con poco valor. Entonces F retoma la palabra ‘loco’ para acentuar la ambigüedad que ofrece el contexto entre ‘frenético’ y ‘auténticamente loco’. Al decir “como lo que era”, F reafirma la evaluación de S, recordando que, sin lugar a dudas, se trataba de un verdadero loco.

## II.2. Información muy confiable

En (8) la hablante S, que es profesora de francés en un instituto de La Plata, explica que ella vive ahora en Buenos Aires y que debe viajar cada semana para trabajar. S es platense, pero se ha acostumbrado tanto a la vida de Buenos Aires que cada vez se le hace psicológicamente más difícil tener que viajar mientras continúa viviendo en La Plata. La norma para los platenses es querer vivir o volver a su ciudad natal tan querida; por eso, S quiere convencer a F de que su caso es distinto:

- (8) F: o sea que - ¿hoy mismo ya venías de Buenos Aires - hoy?  
 [ ]  
 S: claro, sí  
 F: a:h yo creía que te que venías el dom cuando me dijiste la otra vez que fuiste a Buenos Aires, yo creí -a que domingo, una cosa así:  
 S: no no no no no:, yo me instalo allá y: este y vengo los martes a - antes venía los: e lunes a la noche () ahora vengo los [riéndose] martes a la mañana  
 A: a:h  
 S: y ya no tengo ganas de: este, **te digo** ya el lunes a la noche estoy este: - pensando que tengo ganas de quedarme – allá

El oyente debe creer que lo que dice el hablante es verdad. Este uso tiene relación con la escala de confiabilidad. El hablante siente que su información

<sup>4</sup> En inglés: ‘homeless’.

es percibida como poco confiable, y entonces intenta dar a esa información un estatus mucho más alto en la escala de confiabilidad. En otras palabras, el hablante trata de convencer al oyente de que no está exagerando. Así, *te digo* está teñido de un matiz deóntico: es casi una orden dada al oyente de que *debe* creer lo dicho por el hablante; equivale a decir *créeme*. En este uso, *digo* siempre aparece junto al objeto indirecto, o sea, a quien va dirigido el mensaje.

Por otra parte, (8) es también un ejemplo de *te digo* usado para contrastar la información presentada con las expectativas dentro de la cultura platense, según se verá más adelante en **III.2**.

Finalmente, los ejemplos (9-12) registran los usos correspondientes a **III**:

### **III. CONTRASTE ENTRE LA INFORMACIÓN DADA Y LOS RECURSOS VERBALES, O LAS EXPECTATIVAS**

Hay casos en que el hablante percibe e indica que hay un desajuste entre *el conocimiento* y *los recursos verbales* que utiliza. Así, puede contrastar la información presentada con las categorías semánticas elegidas para comunicar esa información. Si busca la palabra exacta para expresar lo que desea, o no está de acuerdo con una palabra que otro hablante ha usado, puede encontrar indicadores evidenciales para señalar la disparidad entre lo expresado y lo que se intenta expresar:

#### **III.1. Contraste entre el conocimiento de la información y los recursos verbales**

El ejemplo (6), presentado más arriba en la sección **II.1. Información poco confiable**, ilustra también **III.1**. Se repite a continuación con el número (9) para facilitar la lectura:

- (9) L: leía o::h leía ta:nto () y: bueno, todos lo: - él casi era como una digamos ¿cómo se llama ese lugar donde van todos los este: () los y digamos los inmigrantes que se van a: a asesorar de cosas? allá en casa era: () una ¿cómo **te diría?** u - una embajada no, ya es muy - muy grande muy grande decir [...]

M: una embajada

L: ¿una qué?

A: [...] un consulado

F: un consulado

L: un consulado una cosa así, porque todos los p:apeles se los hacía papá.

En (9) L está hablando de que su padre, un inmigrante italiano que tenía un negocio de almacén, ayudaba a otros inmigrantes que no sabían leer ni escribir. La idea que quiere dar es que el almacén funcionaba como un consulado “porque todos los papeles se los hacía papá”. L propone la palabra ‘embajada’ para completar su explicación, pero sabe que no representa el concepto de lo que intenta decir. Son otros participantes quienes dan una palabra más aproximada: ‘consulado’.

En este ejemplo específico el indicador evidencial es toda la pregunta introspectiva “¿cómo *te diría*?”. En los datos consultados también hay ejemplos con el uso de *te diría/no te diría* + la categoría propuesta, que no se incluyen en este trabajo por razones de espacio.

En (10) L corrige a M en relación a la palabra *inmediatamente*:

(10) M: pero vamos a comer con ellos inmediatamente.

L: **no te digo** inmediatamente.

La hablante L utiliza el indicador evidencial *no te digo* para aclarar que la forma en que M ha usado *inmediatamente* no es apropiada; de este modo señala que en ese contexto ‘inmediatamente’ no significa ‘inmediatamente’ en forma literal. En otros ejemplos similares del rioplatense se encuentra la expresión *no te diría* con el mismo valor.

Por último, en rioplatense se usan los indicadores evidenciales *te diré* y *te digo* con el propósito de contrastar la información comunicada con la información esperada. Esta función pragmática de *contrario a lo esperado* puede darse por expectativas en relación a lo que es la norma para el hablante y los otros miembros de la misma cultura, o por expectativas producidas por el contexto dado:

### III.2. Contraste con expectativas

La expresión *te diré* se usa en (11) para marcar que la información dada está en contraste con las expectativas derivadas del contexto situacional:

(11) PA: No fuimos al museo, no fuimos a la bodega, no fuimos a la parte histórica... Nos perdimos un montón de cosas **te diré**.

La hablante PA está contrariada con lo que ha pasado en sus vacaciones. Refiere que ese grupo de personas esperaba algo que no sucedió.

En (12) el hablante C relata sobre el comportamiento insólito de los enfermeros de un hospital, quienes competían en el solárium, dentro del edificio, cuando no tenían pacientes y estaban aburridos. C explica que esta habitación era inusualmente grande, y que precisamente sus dimensiones hacían posible la corrida de carreras con pequeñas motocicletas:

(12) C: (y se sentaban) - entraban los muchachos cuando se aburrían - si no hacían carreras, se subían - a las bambinas y corrían adentro del solarium.

F: ¿las qué?

C: las bambin - las motitos así tra: - las chiquiti:tas, como la que tenía:: Patricia- y corrían adentro del solárium. El solárium es una cosa gra:nde. **Te diré** que tendrá (0.6) y:: debe tener como: quince me:tros porr ci:nco, es una co:sa - y ahí - entonces alrededor de las sillas, viste, y los monos<sup>5</sup> ahí apostando. [risas]

Los oyentes del relato se sorprenden de que se pudieran correr carreras con motos en el interior de un edificio, por eso C aclara que, contrario a lo esperado (*te diré*), el solárium tendría unos quince metros de largo.

De acuerdo al análisis de los ejemplos presentados en esta sección se puede observar que las distintas funciones evidenciales de *decir* resultan del contenido semántico del verbo, suficientemente amplio para permitir alojar tales funciones sin contradecir su significado básico de *hablar y expresar*. En lo que se dice hay una creencia, una opinión o una inferencia. Se puede comunicar lo que se expresa en un escrito, lo que hablaron otros, o la voz del pensamiento. Se puede decir algo convenciendo, o decir con poco convencimiento en lo que se dice; y también, decir reconociendo que lo que decimos puede causar sorpresa.

#### 4. DICE: DE VERBO A PARTÍCULA DISCURSIVA EVIDENCIAL

Como fuera expresado anteriormente, en español la evidencialidad no se presenta codificada en morfemas específicos. En los datos analizados, por otra parte, se puede observar el caso especial de la forma de tercera persona

<sup>5</sup> En lengua muy informal *monos* significa *tipos*.

singular del presente de indicativo del verbo *decir*, que se encuentra en un proceso de cambio a marca evidencial. Así, esta sección se centra en los usos de la forma *dice* que indican un movimiento de verbo a partícula discursiva evidencial para marcar que la fuente de conocimiento de la información es el lenguaje oral.

El cambio está en progreso: la marca evidencial *dice* no se usa obligatoriamente, sino que alterna con los usos tradicionales de *decir* como verbo propiamente dicho cuando se necesita indicar que la información fue obtenida *de oídas*. En los siguientes ejemplos se observan usos que pueden interpretarse como evidencia del proceso de cambio referido.

A veces los hablantes insertan abundantes marcas indicadoras de que el lenguaje oral ha sido la fuente de la información que proveen, como en (13-14):

- (13) A: y entonces me **dice**: no: - le **dice** - le **digo** me quería comprar le **digo** para llevarle a mi mamá le **digo** a Ana. (0.4) o:y **dice** y pensar que yo le podía - este por qué no les mandaba por correo, yo no tenía el cupón. Porque si no, eran muy caras ¿no? muy muy caras (0.6) entonces **dice**: entonces yo creo que tengo algunas, entonces la man - le - llamó a Fred por teléfono, y le **dijo** que le trajera - la que todavía tenía ella. [...]. Entonces **dice to:me**, **dice**, le traje ésta para que se la lleve a su mamá de mi parte me **dijo**. Te la trajo ella.
- (14) A: es decir, no me **dijo** que: no me **dijo** que no se lo diera me **dijo** que - e me lo mostró, **digo** uy todavía le quedaron le **dije**, tenía bastante o sea que e e: que: - porque le pregunté ¿acá lo venden al final? **dice** no sé porque todavía tengo me **dijo**, finalmente

En (13) y (14) las formas de *decir* conservan la marca de persona gramatical. Por otra parte, la falta de constancia en cuanto a la referencia temporal de pasado (que ocurre en varias instancias) y, sobre todo, la abundante repetición de este verbo, sugieren que en este uso *decir* está en transición hacia la posible pérdida de su condición de verbo. En estos ejemplos aparece en un estadio intermedio, funcionando en parte como verbo, y en parte como marcador evidencial: por un lado, conserva todavía la morfología verbal, pero por otro, se repite en forma periódica para recordar la fuente de la información.

En (15-17), los hablantes utilizan *dice* como marca reiterada del discurso indirecto: ellos optan por repetir la fuente de conocimiento, casi innecesariamente, aunque se refiera al mismo individuo. Este mecanismo sugiere que, en este uso, *dice* funciona más como partícula invariable que como verbo:

- (15) P: ahora le estaba diciendo acá a Norma que había venido ese día [...], **dice**, vio doña Pepa eso que pasó **dice** con - con los televisores **dice**- el que toma nota ¿tiene necesidad de repetirlo en voz alta? **dice** eh
- (16) P: [...] uno o dos días. Entonces **dice** u:h no sé cómo vino la conversación, alguien que vino acá a hablar por teléfono **dice** () ¿no se enteraron ustedes de lo que pasó con la chancha? () entonces le dije ¿la chancha? le digo y no nosotros no [...] no sabíamos nada () entonces **dice**, por allá **dice**, por [la calle] 505 () pasando la fábrica de armas **dice** () había una () una chancha () **dice**, que tenía chanchitos
- (17) P: y **dice** p - él reía como un loco **dice** el chofer **dice**

En estos ejemplos, *dice* está usado como indicador repetido para que el oyente recuerde que la narración no se ha originado en el hablante mismo, quien sólo está citando la fuente de información. Es interesante notar que *dice* es usado a veces entre oraciones, y otras, en forma intra-oracional, como en (16): “pasando la fábrica de armas **dice** () había una () una chancha () **dice**, que tenía chanchitos”.

El ejemplo (17) es parte de la narración de una anécdota que ocurre en un taxi, que hace reír al conductor. Aquí las tres instancias de *dice* tienen el mismo referente: no debe interpretarse que es el chofer quien *dice*, sino la vecina de P, que es la fuente de información que narró la anécdota a la hablante.

En (18), a continuación, *dice* es más una indicación evidencial de lo que pensaba y solía decir el abuelo (reinterpretado por la hablante N) que un uso prototípico del verbo *decir*:

- (18) N: no: no: no: no murió: - murió antes él. Yo me acuerdo, porque cuando llegó el hombre a la luna yo me acuerd - si hubiera estado el abuelo, (0.6) que quería ver el hombre a la luna. Quería ver llegar el hombre a la luna. (0.8) **Dice**, no me quiero morir sin llegar a - sin que llegue el hombre a la luna. (2.4) Y no, y murió antes.

En este contexto, la forma esperada para referirse a un pasado habitual sería el Imperfecto *decía*, y no el Pretérito *dijo*, que es la que usualmente podría ser reemplazada por la de Presente ‘histórico’. En este ejemplo, *dice* presenta pérdida de tiempo y aspecto.

En (19) la primera instancia de *dice* es usada en forma prototípica por la hablante M para indicar discurso indirecto. O sea, advierte a los oyentes que la información provista fue obtenida por medio del lenguaje oral, o *de oídas*. M cita a un médico que habla de la salud en televisión, quien siempre dice que las vitaminas son necesarias pero que conviene tomarlas en pequeñas dosis:

(19) M:no: bueno pero es - pero es un docto:r que enseña de los dolores reumá:ticos (y) **dice** que las vitaminas que son necesarias. [cita las palabras del doctor] Lo que le conviene, cuando se toma vitamina, este: - repártalas en varias tomas. En vez de tomarse una toda una vez (o.8) la reparte en varias tomas

A: a mí me parece que es más efectivo. Y siempre **dice** tomar las vitaminas después de comer. Así se fijan más.

Sin embargo, en la segunda instancia de *dice*, la hablante A no utiliza esa forma con el mismo valor, sino con significado impersonal, equivalente a *se dice*. A está viviendo en los Estados Unidos y no conoce al médico referido que aparece en la televisión argentina. En realidad, A expresa que es habitual oír la recomendación de no ingerir vitaminas con el estómago vacío. Por lo tanto, las dos formas de *dice* incluidas en (19) tienen distintos referentes: en la primera instancia se trata del médico citado; en la segunda es una voz impersonal. En este uso *dice* ha perdido el morfema flexivo *se* (en este contexto, una forma de déictico impersonal). O sea, que *dice* ha perdido la agentividad de un referente específico; ha desplazado su voz activa a una voz media. Y esto, a su vez, tiene influencia en el lexema de *decir*, ya que su significado no se interpreta en forma literal sino más bien como información sabida, como conocimiento de dominio público. De este modo, *dice* pasa a funcionar como marca evidencial más abstracta.

Dentro de la clase **I.3.2.** (cf. Tabla 1) se encuentran ejemplos en que *dice* aparece como forma invariable en proceso de discursivización: funciona como marca evidencial *de oídas* para reportar discurso indirecto. En este uso *dice* no se comporta como verbo sino como una partícula, ya que llega a utilizarse con distintos sujetos, no sólo con la tercera persona singular. Tiene

referencia de pasado y puede significar *digo/dije*, *decís/dijiste*, *decimos/dijimos*, (Uds., ellos) *dicen/dijeron*.

En el contexto de (20), (ellos) *dicen* es la forma esperada para las dos últimas instancias de *decir*, ya que tienen el mismo referente. Pero en la última el hablante utiliza *dice* con la misma referencia de plural: los profesionales y universitarios que no toman conciencia de la importancia de participar en actividades extra curriculares para crecer en la profesión de bioquímico:

(20)C: es decir, si vos te pasás todo el día: - cruzado de bra:zos o: p:ensando en la luna, y: no progresás, bueno, e:s lógico. Pero ya que intelectualmente, no::, (0.2) materialmente, pero bueno, fenómeno. Es - e: - es una cosmovisión muy particular. Pero si vos querés - e::: m - organizarte económicamente en - en cosas (0.6) mirá, lo básico es lo siguiente. Yo lo veo en muchos e de: mis compañeros, y demás que l - les **decís**, mirá, esta noche: los invito a un cu:urso. (0.4) Vengan a un cu:urso, entrada pa:ga, un lunch, e que viene después del curso. Me **dicen** no, pero no tengo qué ponerme. (0.4) **Dice** y p - son profesionales, universitarios, no?

En (21) dos hermanas le cuentan a la hija/sobrina de ellas las maravillas que hace una repostera de televisión, que produce decoraciones muy originales, y a quien consideran una gran artista. La hablante A queda muy interesada por las descripciones, y en un momento dado pregunta por algo que mencionó antes su tía: ciertas muñecas hechas como parte de la decoración. A quiere saber si oyó bien, y le pregunta a su tía si en verdad ha hablado de muñecas: “¿y [también] muñecas, **dice**? ¿muñecas?” En rioplatense, las formas esperadas de *decir* en ese contexto habrían sido *decías*, o posiblemente *decís* o *dijiste*. Sin embargo, A usa en su lugar la forma invariable *dice*:

(21) M:y es una arti:sta [...]  
 A: no? e usa usa () de: elementos de:=  
 M:=hace tortas como si fueran un cuadro, con el reborde y  
 [  
 L: unas muñecas, nena  
 M:todo () y una pareja que está bailando el minué como en  
 aquellas épocas que vos veías esos cuadros anti:guos [...]



y ella dice y es muy fácil. Claro, ella lo enseña y es fácil [...] claro () lo que hace ahí () nunca le sale perfecto () porque le dan unos segu:ndos nada más [...]

L: cuando hace postales

M: cuando hace postales, para las fiestas, l (...) – es [...] - e - una cosa que emociona realmente

A: y muñecas, **dice**? ¿muñecas?

M: muñe:cas, muñecas

[

L: unas muñecas, de ésas con: con: este:: muñecas antiguas, digamos. Damas antiguas. () Pero con la:s pelucas ésas

M: pelucas bla:ncas, así

El ejemplo (22) ocurre en una conversación en que las participantes están discutiendo sobre los distintos nombres de los lactobacilos. La hablante A explica que cuando intentaba comprarlos en España, en principio no le entendían qué necesitaba, hasta que en algunas farmacias llegaron a aclararle que allí se conocían por otro nombre. En este uso, *dice* es una marca evidencial más que un verbo: los que dicen que allí se llaman ‘lactófilus’ son varios farmacéuticos. La forma esperada en (22) sería *decían* (tercera persona plural, pretérito imperfecto del indicativo): cada vez que A preguntaba le decían el verdadero nombre del producto:

(22)A: bueno. Entonces éstos son e son los acidófilus. Que en lo - que en España yo los pedía y se llamaban lactófilus. Sí. Yo les - se los describía. **Dice** sí, pero eso se llama lactófilus ahora.

En (23) *dice* se presenta similar al ejemplo anterior, excepto que su uso como partícula discursiva evidencial se encuentra aumentado por la reiteración. Aquí V, que está embarazada, se refiere a su parto, que ocurrirá en poco tiempo. Sus compañeras la han convencido de que la presencia del padre durante el parto será muy beneficiosa para ella, ya que después de esa experiencia su marido no será capaz de negarle nada. La partícula *dice* tiene el mismo significado de *decían*, usado por V en: “las chicas hoy en el curso me decían”. O sea, la referencia temporal es de pasado, la aspectual sería imperfectiva, y la de persona: tercera plural. La partícula invariable *dice* se repite periódicamente para recordar a los oyentes la fuente *de oídas*:

(23)V: así - y al final, las chicas hoy en el curso me decían, al final va a ser así, Violeta, **dice**. Porque el padre, (0.4) ¿después del parto? te permite cualquier cosa, **dice**. Después de que te vio, viste, parir, **dice**, bueno, pone:le, no sé:, viste, hacía lo que [riéndose] quie:ras. [...] así que - y él va a estar en el parto.

En los ejemplos analizados en esta sección la forma *dice* se va alejando del valor léxico del verbo *decir* para reinterpretarse como marcador evidencial invariable dentro del discurso oral. De este modo, los datos del rioplatense indican que *dice* se encuentra en un proceso de transformación de verbo a partícula evidencial *de oídas*, lo que resulta –según propongo– en un movimiento hacia la discursivización. Es el significado del lexema del verbo *decir* lo que posibilita su cambio semántico a marca de evidencialidad, un concepto más abstracto que el concepto de *decir*, que es más específico. Como partícula, *dice* ha perdido sus referencias de persona, aspecto, tiempo y voz, y también su significado literal de *hablar o expresar* algo. Indirectamente conserva aún su modalidad epistémica, ya que su condición de indicador de información reportada implica algún porcentaje de posibilidad de alejarse de la verdad por no provenir de fuente directa. La partícula discursiva *dice* solamente retiene su valor de deíctico evidencial; y por lo tanto, en este sentido, constituye un índice, en el sentido de Pierce (Atkin 2013).

## 5. MOVIMIENTO HACIA LA DISCURSIVIZACIÓN

Según Heine et al. (1991), el movimiento hacia la gramaticalización utiliza una estrategia cognoscitiva de tipo ‘resolución de problemas’, por la que los conceptos más inmediatos y accesibles a la experiencia humana se emplean para expresar conceptos más abstractos. Ocampo (2006) distingue entre movimientos de gramaticalización y de discursivización, pero la observación de Heine et al. también es aplicable a estos últimos.

En su estudio sobre la partícula discursiva *claro*, Ocampo (2006: 317) habla de cuatro movimientos diacrónicos posibles. En los casos de *gramaticalización* la trayectoria se produce de léxico a gramática; por otra parte, sería *degramaticalización* cuando un elemento cambia de un estatus más gramatical a menos gramatical, o más lexical; luego habría

una tercera posibilidad: *recodificación*, en que el movimiento podría darse de discurso a léxico, o de discurso a gramática; y, por último, *discursivización*, cuando el cambio ocurre de léxico a gramática y luego a discurso o, en un trayecto menor, de gramática a discurso. Autores como Martín Zorroaquiño y Portolés Lázaro (1999) presentan una visión del proceso de cambio consistente con la de Ocampo en relación a marcadores del discurso.

Aplicando las propiedades utilizadas por Ocampo (2006) en el análisis de *claro*, podemos decir que *dice* se encuentra en transición de verbo a partícula discursiva. Por un lado, comparte dos propiedades con los elementos que se gramaticalizan:

- reducción morfológica
- cambio semántico de concreto a abstracto

Por otra parte, son más las propiedades que *dice* tiene en común con los elementos que se discursivizan:

- no pertenece a un sistema (pierde la noción de *valor*; escapa a un paradigma; se mueve fuera de la sintaxis)
- no hay fusión con ningún constituyente
- pasa al plano de la estructura conversacional
- aumenta el *alcance* o radio de influencia<sup>6</sup>
- no hay obligatoriedad de uso (depende de la elección del hablante, según sus propósitos)

Las propiedades listadas, que *dice* comparte con otros elementos que han sufrido un cambio hacia la discursivización, son indicadoras de que esta forma invariable constituye también una partícula que se encuentra en el mismo proceso.

## 6. CONCLUSIONES

Después de referirnos a la noción de evidencialidad y de presentar una taxonomía, hemos anotado las formas de *decir* usadas en rioplatense. En los ejemplos presentados se registran instancias en que *dice* pierde sus características verbales: persona, tiempo, aspecto y voz. En este proceso

---

<sup>6</sup> En inglés: 'scope'.

también se pierde la transitividad de *decir* y, simultáneamente a los cambios gramaticales, el lexema verbal transforma su significación original: de ‘decir algo específico’, se pasa a la idea de ‘reportar una fuente indirecta’, y más, a ‘reportar’ o ‘reportarse’ (en voz media). De este modo, *dice* se convierte en índice evidencial *de oídas*, o sea, en una partícula discursiva cuya única función es indicar información reportada.



## REFERENCES

- Aikhenvald, Alexandra (2004): *Evidentiality*. Oxford: Oxford University Press.
- (2006): “Evidentiality in grammar”, en Keith Brown (ed.): *Encyclopedia of Languages and Linguistics*. Oxford: Elsevier, pp. 320-325.
- Aksu-Koç, Aihan y Dan Slobin (1986): “A psychological account of the development and use of evidentials in Turkish”, en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 159-167.
- Aoki, Harao (1986: “Evidentials in Japanese”, in Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 223-238.
- Atkin, Albert (2013): “Peirce’s Theory of Signs”, en Edward Zalta (ed.): *The Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/>.
- Bermúdez, Fernando (2005): *Evidencialidad: la codificación lingüística del punto de vista*. Tesis Doctoral. Universidad de Estocolmo.
- Boas, Franz (1938): “Language”, en Franz Boas (ed.): *General Anthropology*. Boston/New York: D. C. Heath and Company, pp 124–145.
- Chafe, Wallace (1986): “Evidentiality in English conversation and academic writing”, en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 261-272.
- DeLancey, Scott (1986): “Evidentiality and volitionality in Tibetan”, en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp.203-213.

- Friedman, Victor (1986): "Evidentiality in the Balcans: Bulgarian, Mecedonian, and Albanian", en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 168-187.
- Givón, Talmy (1989): *Mind, Code and Context. Essays in Pragmatics*. Hillsdale, NJ: Lawrence Erlbaum.
- Gordon, Lynn (1986): "The development of evidentials in Maricopa", en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 75-88.
- Hardman, Martha James (1986): "Data-source marking in the Jaqi languages", en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 113-136.
- Heine, Bernd, Ulrike Claudi y Friederike Hünemeyer (1991): *Grammaticalization*. Chicago and London: The University of Chicago Press.
- Jakobson, Roman (1957): *Shifters, verbal categories, and the Russian verb*. Cambridge: Harvard University Press.
- Martín Zorraquino, María Antonia y José Portolés Lázaro (1999): "Los marcadores del discurso", en Ignacio Bosque y Violeta Demonte (eds.): *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, pp. 4051-4213.
- Ocampo, Francisco (2006): "Movement towards discourse is not grammaticalization: The evolution of *claro* from adjective to discourse particle in spoken Spanish", en Nuria Sagarra y Almeida Jacqueline Toribio (eds.): *Selected Proceedings of the 9<sup>th</sup> Hispanic Linguistic Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 308-319.  
<http://www.lingref.com/cpp/hls/9/paper1388.pdf>.
- Papafragou, Anna, Peggi Li, Youngon Choi y Chung-hye Han (2007): "Evidentiality in language and cognition", *Cognition* 103, pp. 253-299.  
<http://dx.doi.org/10.1016/j.cognition.2006.04.001>
- Schlichter, Alice (1986): "The origins and deictic nature of Wintu Evidentials", en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 46-59.

- Speas, Peggy (2008): “On the syntax and semantics of evidentials”, *Language and Linguistics Compass*, 2/5, pp. 940-965.  
<http://dx.doi.org/10.1111/j.1749-818X.2008.00069.x>
- Weber, David (1986): “Information perspective, profile, and pattern in Quechua”, en Wallace Chafe y Johanna Nichols (eds.): *Evidentiality: The Coding of Epistemology in Language*. Norwood, NJ: Ablex, pp. 137-155.





## Apéndice

### CONVENCIONES UTILIZADAS EN LA TRANSCRIPCIÓN

- Corchete “[” entre dos emisiones marca superposición.
- “=” indica superposición parcial entre turnos.
- “-”: hiato en el habla.
- Un silencio breve se indica con paréntesis “()”.
- Los números entre paréntesis “(1.4)” representan un silencio en segundos y décimas.
- Los subrayados indican prominencia de intensidad o tono agudo.
- “:” marcan duración en la pronunciación de vocales o consonantes.
- “[...]” indican que se omite parte de la transcripción.
- Cuando no se puede asegurar la precisión de la transcripción el material aparece entre paréntesis.
- Tres puntos entre paréntesis indican que esa parte de la emisión es ininteligible.
- Información adicional/palabras agregadas para mejorar la comprensión aparecen entre corchetes “[se ríe]”.
- En cada ejemplo, lo que es relevante para el análisis aparece en **negrilla**.
- Los nombres de los participantes han sido cambiados para proteger su identidad.





## **Excess, Deficiency, and Mean: Resumptive Pronouns, Prepositional Phrase Chopping, and Gaps in Spanish Oblique Relative Clauses**

Álvaro Cerrón-Palomino  
*Arizona State University*

**Abstract:** This study is a variationist account of a three-way alternation found in Spanish oblique relative clauses (RCs): gaps, resumptive pronouns (RPs), and prepositional phrase (PP) chopping. Using four data sets of the variety spoken in Lima, Peru, three separate multivariate analyses were run, showing that each of the variants under scrutiny corresponds to a different set of constraints. First, gaps are favored by RCs containing only verbs, inanimate antecedents, and plural nouns. Secondly, RPs are promoted by animate antecedents, semantically full prepositions, stacked RCs, and indefinite antecedents. Finally, PP-chopping is favored by inanimate and singular antecedents. This article shows, on the one hand, the relevance of animacy in the choice of oblique RC strategies in Spanish. On the other hand, this paper suggests that the variants studied are not affected by the syntactic complexity or simplicity of the structure antecedent + RC, which were thought to correlate, respectively, with difficulty and ease of processing.

**Keywords:** Relative Clauses, Resumption, Preposition Chopping.



## 1. INTRODUCTION

The alternation between resumptive pronouns (RPs) in Spanish relative clauses (RCs) headed by the complementizer *que* (that) like (1), and gap-containing or pied-piping RCs like (2), has received particular attention from variationist studies in the last three decades (Herrera 1990; Silva-Corvalán 1996; Bentivoglio 2003; Samper *et al.* 2005; Cerrón-Palomino 2006).

- (1) Con una amiga, con Lilian B., con una amiga<sub>i</sub> que hace tiempo no vuelo con **ella**<sub>i</sub>.  
'With a friend, with Lilian B., with a friend<sub>i</sub> that I haven't flown with **her**<sub>i</sub> in a while'.

(D2,11N,F34)<sup>1</sup>

- (2) Con una amiga, con Lilian B., con una amiga<sub>i</sub> con (la) que hace tiempo no vuelo  $\emptyset$ <sub>i</sub>.  
'With a friend, with Lilian B., with a friend<sub>i</sub> with which I haven't flown  $\emptyset$ <sub>i</sub> in a while'.

The aforementioned studies aimed at determining the constraints governing the presence of RPs in subject, direct object and oblique RCs<sup>2</sup>, in a number of Spanish varieties. Interestingly, oblique RCs, like the ones presented in (1)-(2), offer a more complex variation context in Spanish: speakers have a third competing structure at their disposal, known as prepositional phrase (PP) chopping<sup>3</sup>, which is illustrated in (3).

- (3) Con una amiga, con Lilian B., con una amiga<sub>i</sub> ~~con~~ que hace tiempo no vuelo  $\emptyset$ <sub>i</sub>.  
'With a friend, with Lilian B., with a friend<sub>i</sub> ~~with~~ which I haven't flown  $\emptyset$ <sub>i</sub> in a while'.

However, in contrast to the research accounting for RP occurrence, there is a lack of studies explaining the presence of PP-chopping in Spanish oblique RCs. In order to fill this gap, the present study aims at explaining the

---

<sup>1</sup> Coding of the speaker that produced the utterance. *D2* refers to Data Set 2, *11* is the number of the speaker of that data set, *N* stands for not college-educated, *F* indicates female gender, and *34* is the age of the speaker at the time of the interview. When no source is mentioned, like utterances (2)-(3), it means that it is an example produced by the researcher.

<sup>2</sup> Indirect object RCs were excluded from the quantificational analyses because of their categorical display of RPs.

<sup>3</sup> Tarallo (1986).

linguistic constraints that govern each one of the three choices oblique RCs offer: RPs, PP-chopping, and gaps.

## 2. PREVIOUS STUDIES

As mentioned in the preceding section, no studies of PP-chopping in Spanish oblique RCs have been conducted before. The only PP-chopping study available for an Iberian Peninsula language is Tarallo's (1986) variationist account of the RP/PP-chopping alternation in RCs headed by the complementizer *que* in spoken Brazilian Portuguese<sup>4</sup>.

The aforementioned three-way variation is presented in examples (4)-(7). On the one hand, we have the alternation between the RP-containing RC and the standard gap-containing RC shown in (4)-(5); on the other hand, in (6)-(7) we can see the alternation between the PP-chopping RC and the gap-containing RC<sup>5</sup>.

- (4) Estava, lembra, con aquela sacolinha, que você ia na faculdade com **ela**.  
'I had, remember, that small bag, that you went to school with **it**'.
- (5) Estava, lembra, con aquela sacolinha, com que / com a qual você ia na faculdade  $\emptyset$ .  
'I had, remember, that small bag, with which you went to school  $\emptyset$ '.
- (6) Eu não ligo para piscina. É uma coisa, que eu não ligo para  $\emptyset$ .  
'I don't care for a pool. It is something, that I don't care for  $\emptyset$ '.
- (7) Eu não ligo para piscina. É uma coisa, para a qual eu não ligo  $\emptyset$ .  
'I don't care for a pool. It is something, for which I don't care  $\emptyset$ '.

The distribution of the alternating variants is shown in Table 1 (Adapted from Tarallo 1986: 259). First, it is evident that PP-chopping is the preferred variant in Brazilian Portuguese, accounting for 73.8% of the total number of cases of RCs analyzed. Secondly, while RPs are the second most common strategy, at a rate of 15.2% of the whole data set, they are far from being a

<sup>4</sup> Tarallo's analysis is not restricted exclusively to oblique RCs, but it is also extended to indirect object and genitive RCs, which can be introduced by prepositional phrases.

<sup>5</sup> All these examples are taken from Tarallo (1986: 249).

popular structure in the speech community studied. Finally, gap-containing RCs represent only 6.5% of the RCs studied. This frequency distribution indicates that standard gap-containing RCs are losing ground to both non-standard RC structures, in particular, to PP-chopping.

Gap		PP-Chopping		RP	
%	N	%	N	%	N
6.5	21	78.3	254	15.2	49
Total N		324			

**Table 1.** Distribution of RC variants with a prepositional phrase in Brazilian Portuguese

In regards to the multivariate analysis he conducted using VARBRUL, Tarallo found that a subset of contexts that favor the occurrence of RPs are “structures which are more loosely woven into discourse” (1986: 253), such as non-restrictive RCs, material before *que* and after *que*, the greater distance between the head of the antecedent noun phrase (NP) and the ‘would be’ gap, and the position of the relative with regard to the matrix. Thus, he concludes that the main function of RPs in RCs is to tighten the relationship between the antecedent NP and the RC.

However, some factors of a different nature were also found to be statistically significant, such as semantic features of the antecedent NP head: NPs with human referents and indefinite NPs. Additionally, the genitive syntactic function of an RC, existential structures, and RCs located at the right position of the main clause promote the use of RPs.

Although he does not offer a separate analysis for PP-chopping, Tarallo suggests that this strategy stands in complementary distribution with the RP strategy, because some contexts that favor the former variant disfavor the latter one, and vice versa. In particular, the factors he found to increase the number of PP-chopping instances are restrictive RCs, non-human NP heads, and definite NP heads.

Another detail that Tarallo points out is that the PP-chopping strategy does not apply to all of the prepositions in his data. On the one hand, prepositions like *a/para* (for), *com* (with), *de* (of), *de/sobre* (about), *em* (at, in, on), *por* (for, by), and *sobre* (on) were chopped; on the other, prepositional locutions like *em cima* (above), *em baixo* (under), *a respeito* ‘about’ (which



he calls ‘strong’) were stranded. However, because of this almost categorical distribution, Tarallo did not include the preposition type as a factor in the multivariate analysis.

In general, Tarallo’s main claim is that PP-chopping is mostly found in simple syntactic contexts, where the loss of the preposition information does not hamper the parsing of the structure. With regard to RPs, Tarallo concludes that they are used in more complex syntactic environments “that inhibit the filling of the gap” (1986: 252), therefore as a means of recovering the function of the antecedent in the RC.

When we turn to sociolinguistic studies of RPs in Spanish RCs headed by the complementizer *que*, it is important to point out that they do not focus exclusively on oblique RCs, but rather on RPs across different syntactic functions: subject, direct object (DO), and oblique altogether.

One of the earliest sociolinguistic accounts of this variable is Juana Herrera’s (1990) analysis using data from Santa Cruz, Canary Islands. Despite the fact that the number of RP-containing RCs is extremely scarce to carry out a probabilistic analysis (only 25 cases), most of her findings have been supported by studies comprising larger data sets. She found that the linguistic factors favoring RP use were subject and DO RCs, as well as sentential antecedents, and non-restrictive RCs. Examples of these structures are presented below (Herrera Santana 1990: 550-551).

- (8) Pues engancha la rueda de otro vehículo<sub>i</sub>, que **esto**<sub>i</sub> es... que **eso**<sub>i</sub> es pa’ uno reírse.  
 ‘Well, it attaches to the wheel of another vehicle<sub>i</sub>, which **this**<sub>i</sub> is... which **that**<sub>i</sub> is something to laugh about’.
- (9) Creo que es una decisión<sub>i</sub> que no **la**<sub>i</sub> tiene que adoptar ni la Iglesia ni los hombres.  
 ‘I think it is a decision<sub>i</sub> that neither the Church nor men have to adopt **it**<sub>i</sub>’.
- (10) Porque realmente la selectividad la pasa cualquier alumno que haya pasado C.O.U. y que tenga unos conocimientos mínimos<sub>i</sub>, que, por supuesto, **los**<sub>i</sub> tiene que tener.  
 ‘Because, in reality, selection can be achieved by any student that has passed C.O.U. and that has minimal knowledge<sub>i</sub>, which, of course, they have to have **it**<sub>i</sub>’.

First, as we can see, in (8), the pronouns *esto* and then *eso* refer to the sentential structure underlined, signaling its subject function in the RC.

Secondly, in (9) the antecedent *una decisión* is referred to by the DO clitic *la*, which reflects its syntactic function within the RC. Finally, in (10) the clitic *los* is used to indicate the DO function of its antecedent inside the non-restrictive RC.

On the other hand, the social factors she found to promote RP presence were male speakers and educated speakers. Although male speakers favored the non-standard variant, the fact that educated speakers also prefer it suggests that RPs are not a stigmatized variant in Santa Cruz.

The first study featuring a multivariate analysis of RPs in Spanish RCs was Silva-Corvalán's (1996) account of RPs in the Chilean variety of Santiago, the capital. Her results after using VARBRUL ranked proper noun head antecedents<sup>6</sup> and indefinite antecedents as the most favorable context for RPs, followed by DO RCs, non-restrictive RCs, animate antecedents, and material intervening between the antecedent and the RC. Examples of some of these contexts are presented below (Silva-Corvalán 1996: 385-394).

- (11) Esteban<sub>i</sub>, que como cualquier otra persona que no se ha operado, **él**<sub>i</sub> no sabía lo que era una operación, él se sentía inútil.  
 'Esteban<sub>i</sub>, that like any other person who has not been operated on, **he**<sub>i</sub> didn't know what an operation was, he felt helpless'.
- (12) Después le dije yo que era una persona<sub>i</sub> que no valía la pena conversar con **él**<sub>i</sub>.  
 'Afterwards I told him that he was a person<sub>i</sub> that it wasn't worth talking with **him**<sub>i</sub>'.
- (13) Las agüitas que me dio la madre Teodosia, que me **las**<sub>i</sub> mandó con la F. y la I.  
 'The herb teas that Mother Teodosia gave me, that she sent **them**<sub>i</sub> to me with F. and I'.

Example (11) shows the proper noun *Esteban* and its co-referential subject RP *él*. In (12), we can see the animate antecedent effect: *persona* is being resumed by the pronoun *él* as well. (13) is an instance of material –another RC– intervening between the antecedent head *agüitas* and the second RC. According to Silva-Corvalán, the RP in (13) eases the identification of the antecedent's syntactic function in the second RC.

<sup>6</sup> Silva-Corvalán's study was the only one that showed proper noun antecedents as a factor contributing to the occurrence of RPs in Spanish.

Although she doesn't include it in the multivariate analysis, Silva-Corvalán suggests that the weight of the information borne by the antecedent and by the RC plays a role in the use of RPs: the greater the weight they have (the number of complements/adjuncts), the greater the chance for RPs to occur. The rationale behind this processing hypothesis is that RPs are used to facilitate the task of parsing rather complex structures.

With regard to social factors, the only one analyzed in Silva-Corvalán's study, education, yielded no statistically significant results, a fact that supports the impression that RPs are not a stigmatized structure, since they are equally used by both educated and non-educated speakers.

In her study of RP and gap alternation in RCs in Caracas, Venezuela, Paola Bentivoglio (2003) followed Silva-Corvalán's methodology for the most part. In addition, she included other social factors in her analysis, such as age, gender, and socioeconomic status.

Her results, after entering the data in VARBRUL, show that the factors favoring RP use are DO and oblique RCs, followed by non-restrictive RCs, indefinite antecedents, upper class speakers, material intervening between the antecedent and *que* and, lastly, animate antecedents.

The only social factor that proved to be significant in her study was socioeconomic status: the highest group favors the RP variant significantly more than the middle and lower groups. This result makes Bentivoglio conclude that RPs are, indeed, a stigma-free variant.

In 2005, Samper, Hernández and Pérez, based on the aforementioned study of Silva-Corvalán (1996), analyzed the linguistic factors constraining RP occurrence in the speech of La Palma de Gran Canaria, although no probabilistic analysis was carried out.

The constraints found to be significant in promoting RP presence in this study are non-restrictive RCs, indefinite antecedents, and material intervening between the antecedent and *que*. However, DO and oblique RPs had almost an identical distribution of frequencies, as opposed to the greater frequency of DO RPs over oblique RPs in Chile and Venezuela. Also, animate antecedents did not favor RPs in a statistically significant way when compared to inanimate ones. In addition, a factor that was proposed for the first time, *i.e.* RCs in which the distance between *que* and the would-be gap is greater, proved to be a favorable context for RP occurrence.

Cerrón-Palomino's (2006) study focused on the internal factors constraining the occurrence of RPs in the Spanish of Lima, Peru. His results

after using Goldvarb X were similar to the ones outlined above, although the ranking of the factors in the multivariate analysis was different. The constraints favoring RP occurrence were, in order of probabilistic strength, DO RCs, non-restrictive RCs, RCs with material intervening between the complementizer *que* and the subordinate verb, indefinite antecedents, and animate antecedents.

In sum, the results obtained for the overall occurrence of RPs show a very similar pattern across the speech communities studied. The main dialectal differences between these communities lie in the ranking of the factors found to condition RP use.

### 3. DATA AND METHODOLOGY

In order to analyze the aforementioned three-way variation attested in Spanish oblique RCs<sup>7</sup>, four data sets containing utterances of *Limeño* speakers were utilized in this study:

1. Transcriptions of recorded speech of 18 speakers<sup>8</sup> of the *Limeño* variety. (D1)
2. Jotted-down utterances of 11 of the aforementioned 18 speakers, in addition to the recording sessions. (D2)
3. Transcriptions of interviews of 14 educated speakers of the *Limeño* variety (Caravedo 1989). (D3)
4. Transcriptions of 12 hours of 2006 sessions of the Peruvian Congress<sup>9</sup>. (D4)

Even after using these four data sets, the number of tokens for the present study was very small: only 152 cases of oblique RCs were attested (compared to 1210 subject RCs and 705 DO RCs), a fact that underscores the scarcity of oblique RCs in spoken Peruvian Spanish<sup>10</sup>.

<sup>7</sup> In contrast with Tarallo's (1986) study, which included RCs with PPs belonging to indirect object and genitive RCs, the present one is restricted to oblique RCs only.

<sup>8</sup> 10 speakers were college-educated and 8 were not college-educated.

<sup>9</sup> The Peruvian Congress gathers members from all over the country, but only utterances of the *Limeño* Spanish variety were used in this study.

<sup>10</sup> In an experiment dealing with time-pressure effects in the production of RCs in different Spanish varieties, Cerrón-Palomino (2011) found that speakers often resorted to subject RCs, DO RCs, and even simple sentences when the target structures were oblique RCs. These results corroborate how infrequent the structures under study are in most Spanish varieties.

As it can be inferred from the nature of the data, no social factors such as age, gender, social class or style were controlled for. In consequence, only linguistic factors were analyzed in this study.

A variationist methodology was followed in this study, and the statistical program used to run three separate multivariate analyses was Goldvarb X.

#### 4. FACTORS CONSIDERED IN THIS STUDY

In this section, I will present the factors included in this study, and the predictions expected for the occurrence of RP on the one hand, and PP-chopping on the other, based on the previous studies reviewed in section 2. The predictions for the factors proposed for the first time will also be included below. The main hypothesis in this study is that there is a complementary distribution relationship between RPs and PP-chopping in Peruvian Spanish, along the lines of Tarallo's (1986) findings for Brazilian Portuguese. Since no study has overtly accounted for gaps in Spanish RCs, no predictions were made regarding the constraints of this variant.

In accordance with previous studies, I hypothesize that non-restrictive RCs should favor RPs due to their loose connection with their antecedent, whereas restrictive RCs should favor the occurrence of PP-chopping due to the tight relationship between the antecedent and the RC, which makes the oblique's function recoverability easy. Also, indefinite antecedents are expected to promote the use of RPs: according to Trujillo (1990), the former are undetermined and the latter add determination to them within the RC. In contrast, definite antecedents are expected to favor the presence of PP-chopping. In like manner, animate NP antecedents are estimated to promote RP usage, whereas inanimate antecedents are assumed to favor PP-chopping.

Following Cerrón-Palomino (2006), I hypothesize that "heavy" RCs with material intervening between *que* and the verb (material to the left of the verb) as in (14) should promote RP use, because this material interferes with the processing of the antecedent's syntactic function inside the RC. On the contrary, light RCs, *i.e.* the ones with no material inside except for the verb, as in (15), are supposed to favor PP-chopping, due to their ease of processing and oblique function recoverability.

- (14) Con una amiga, con Lilian B., con una amiga<sub>i</sub> que hace tiempo no vuelo con **ella**<sub>i</sub>.  
 ‘With a friend, with Lilian B., with a friend<sub>i</sub> that I haven’t flown with **her**<sub>i</sub> in a while’.  
 (D2,11N,F34)
- (15) Vamos mañana al restaurante<sub>i</sub> al que querías ir Ø<sub>i</sub>.  
 ‘Let’s go tomorrow to the restaurant<sub>i</sub> to which you wanted to go Ø<sub>i</sub>’.  
 (D2,18N,F28)

Another factor analyzed in this study is the stacking (or lack thereof) of RCs. The second, third or subsequent RC in a series of RCs like the coordinated ones in (16) is expected to favor an RP due to the processing difficulty that the in-between RC(s) could cause. PPs are anticipated to be deleted in RCs that are not coordinated or juxtaposed, as in (17).

- (16) Los dos niños<sub>i</sub> que tardaron en venir, como te digo, pero que... ya, digamos, con **ellos**<sub>i</sub> ha sido difícil viajar.  
 ‘The two kids<sub>i</sub> that took long to come, as I tell you, but that... say, with **them**<sub>i</sub> it has been hard to travel’.  
 (D3,10C,F40)
- (17) Está viviendo con el chico<sub>i</sub> con que se casó Ø<sub>i</sub>.  
 ‘She is living with the guy<sub>i</sub> to which she got married Ø<sub>i</sub>’.  
 (D2,12N,F34)

Building on Tarallo’s observation that certain prepositions favor PP-chopping more than others, I used Alcina and Blecua’s (1991: 835) semantic classification of Spanish prepositions. ‘Full’ prepositions are the ones that have a prominent lexical meaning and have limited distribution, such as *ante* (before), *bajo* (under), *contra* (against), *desde* (from), *entre* (between), *hacia* (toward), *hasta* (until), *para* (for, to), *por* (for, to), *según* (according to), *sin* (without), *sobre* (over), and *tras* (after). In contrast, ‘empty’ prepositions have multiple combination possibilities, and may be used as mere relational marks whose meaning depends on the term they introduce, like *a* (to, for), *con* (with), *de* (of), and *en* (in, on, at).

The hypothesis that I propose is that full prepositions will favor more RPs, as in (18) and (19), whereas empty prepositions will be more prone to be chopped, as in (20) and (21).

- (18) Los noisers, pe' [sic], los noisers<sub>i</sub> que para **ellos**<sub>i</sub> no hay diferencia.  
 'The noisers, the noisers<sub>i</sub> that for **them**<sub>i</sub> there is no difference'.  
 (D2,04C,M33)
- (19) Por ejemplo, Pedro L.<sub>p</sub> que muchas morían por **él**<sub>p</sub>, nunca me gustó.  
 'For instance, Pedro L.<sub>p</sub> that all the girls were crazy for **him**<sub>p</sub>, I never liked'.  
 (D2,13C,F33)
- (20) Hay una revista<sub>i</sub> a-la que yo me afilié Ø<sub>p</sub> que se llama American Baby.  
 'There is a magazine<sub>i</sub> that I subscribed to Ø<sub>p</sub> that is called American Baby'.  
 (D1,14C,F31)
- (21) Yo lo pongo en colegios<sub>i</sub> en que hay pocos niños Ø<sub>i</sub>.  
 'I put him in schools<sub>i</sub> in which there are few children Ø<sub>i</sub>'.  
 (D2,12N,F34)

A new hypothesis proposed in this study, related to Tarallo's (1986) idea about loosely and tightly connected structures, refers to the category selecting the prepositional phrase. It states that RCs selected by a preposition rather than by a verb are a more favorable context for RPs to occur, due to their relative independence with regard to the verb, as in (22). On the other hand, PP-chopping should occur more when the preposition deleted introduces a verb complement rather than an adjunct, because of its tighter relation with the antecedent, as exemplified in (17), repeated as (23).

- (22) Es el tipo de chica<sub>i</sub> que tú quieres hacer algo con **ella**<sub>i</sub>; y ella, no.  
 'She is the type of girl that you want to do something with with **her**<sub>p</sub>, and she does not'.  
 (D2,12N,F34)
- (23) Está viviendo con el chico<sub>i</sub> con-el que se casó Ø<sub>i</sub>.  
 'She is living with the guy<sub>i</sub> to which she got married Ø<sub>i</sub>'.  
 (D2,12N,F34)

The next factor included in the study is whether the RC is embedded in another clause or not. RCs embedded in other subordinate clauses, as (24), due to the processing complexity they pose, should favor RP occurrence.

In contrast, non-embedded RCs, as (25), are thought to promote the occurrence of PP-chopping.

(24) Es el... lo que quieren también los este... los jub... pensionistas,  
que hemos tenido reuniones con **ellos**<sub>i</sub>, han tenido una  
huelga, inclusive.

‘This is the... what the retir.. the pensionists<sub>i</sub> that we have  
had meetings with **them**<sub>i</sub>, have also had a strike’.

(D4,76C,F53)

(25) De repente hay temas<sub>i</sub> ~~sobre~~ los que nuestra propia bancada  
va a requerir nuestra... nuestro voto Ø<sub>i</sub>.

‘Maybe there are matters<sub>i</sub> ~~about~~ which our own party will  
require our... our vote Ø<sub>i</sub>’.

(D4,3C,M45)

Following Tarallo’s (1986) results, I incorporated the position of the RC relative to the main clause. RCs located at the right side of the main clause, as in (26), are expected to favor RPs, whereas RCs located at the center of the main clause as in (27) are thought of as promoting PP-chopping.

(26) Es una persona<sub>i</sub> que yo sí confío en **ella**<sub>i</sub>.

‘It is a person<sub>i</sub> that I do confide in **her**<sub>i</sub>’.

(D2,11N,F34)

(27) Por ese lugar<sub>i</sub> ~~en~~ que hay un barco Ø<sub>i</sub>, en Tomás Marsano.

‘That place<sub>i</sub> ~~in~~ which there is a ship Ø<sub>i</sub>, on Tomás Marsano  
[Avenue]’.

(D2,12N,F34)

Another factor that Tarallo (1986) analyzed and I include in this study is the semantics of the main clause where the RC is hosted. If Peruvian Spanish behaves like Brazilian Portuguese, RCs whose main verb is existential as in (28) should favor RPs. On the other hand, RCs whose main verb is of a different nature, as in (29), are expected to increase the number of PP-chopping instances.

(28) Había una chiquilla, Cristina<sub>i</sub>, que todo el mundo estaba  
detrás de **ella**<sub>i</sub>.

‘There was a girl, Cristina<sub>i</sub>, that everybody was after **her**<sub>i</sub>’.

(D2,7N,M31)



- (29) Lo puse en un colegio<sub>i</sub> ~~en~~ que había nueve niños Ø<sub>i</sub>.  
 ‘I put him in a school<sub>i</sub> ~~in~~ which there were nine children Ø<sub>i</sub>’.  
 (D2,12N,F34)

The last factor included is the grammatical number of the antecedent NP head. Since plural nouns like the antecedent head in (30) are considered to be marked<sup>11</sup> compared to singular ones like the antecedent in (31) and therefore, more cognitively salient, I hypothesize that plural antecedents will be emphasized by the presence of RPs, whereas PPs including singular antecedents will be more easily chopped.

- (30) Hay muchos problemas<sub>i</sub> que nadie se preocupa por **eso**<sub>i</sub>.  
 ‘There are many problems<sub>i</sub> that nobody cares about **it**<sub>i</sub>’.  
 (D2,11N,F34)

- (31) Si yo tengo una región<sub>i</sub> ~~en la~~ que no hay muertes maternas en los últimos dos años Ø<sub>i</sub>, significa que algo está bien.  
 ‘If I have a region<sub>i</sub> ~~in~~ which there are no maternal deaths in the last two years Ø<sub>i</sub>, it means that something is working’.  
 (D4,67C,M52)

**5. RESULTS AND DISCUSSION**

The frequency distribution of these different strategies in Spanish is presented in Table 2, where we see that the most predominant strategy is the standard gap-containing one, followed by PP-chopping, and finally, by RPs.

Gap		PP-Chopping		RP	
%	N	%	N	%	N
59.1	79	34.2	52	13.8	21
Total N		152			

**Table 2.** Distribution of oblique RC variants in Peruvian Spanish

<sup>11</sup> Wurzel (1988), Croft (1996).

This distribution is clearly different from the one in Brazilian Portuguese, shown above in Table 1. Whereas in this language the most frequent strategy is PP-chopping, which represents 78.3% of the oblique RCs, in Spanish it is the standard, gap-containing structure the one that is preferred, accounting for 51.9% of all occurrences. In Portuguese, RPs are the second most frequent structure, with 15.2% of occurrences; by contrast, in Spanish, PP-chopping is the second most frequent structure with 34.2% of occurrences. Finally, the standard gap-containing variant is the least favored one in Portuguese with a mere frequency of 6.5%, while RPs are the least frequent variant in Peruvian Spanish, with a rate of 13.8%.

In sum, while in Brazilian Portuguese the standard form is receding in favor of the non-standard forms, in Peruvian Spanish the standard structure is still prevalent, a fact that may indicate that Spanish is more conservative than Portuguese regarding oblique RC production.

As it was mentioned in section 3, three separate multivariate analyses were carried out to establish the constraints governing the occurrence of each one of the variants of the variable studied, and to determine the probabilistic strength of the constraints found to be significant. The results of these analyses are presented in the following lines.

Table 3 displays the results for the factors contributing to the occurrence of RPs. As it can be seen, out of the eleven factors considered in the analysis, there are only four that are strong predictors of RP occurrence. These factors are, in order of probabilistic strength<sup>12</sup>, animate antecedents, full prepositions, stacked RCs, and indefinite antecedents. All the other factors, indicated by square brackets, were disregarded by Goldvarb as statistically not significant.

These results do not coincide for the most part with the previous Spanish RP results, since non-restrictive RCs and material to the left of the RC verb (between *que* and the verb) were eliminated in the multivariate analysis. The fact that “heavy” RCs with material intervening between *que* and the subordinate verb do not favor RPs suggests that oblique RCs behave differently from the other RCs considered in previous studies (subject and DO RCs) with respect to the weight of the clause: “heavy” oblique RCs do not make it harder for the speaker to keep track of the syntactic function being relativized.

<sup>12</sup> Factor weights with values of .5 and higher favor the variant reported in the table, factor weights with lower values disfavor it. The range is the factor group's relative strength: the higher its value, the better predictors the constraints are.

However, since stacked RCs do favor the occurrence of RPs, it can be assumed that the real processing difficulty that is posed to the speaker arises when a sequence of RCs shares a single antecedent. Inside the RC at stake, it would be hard to keep track of the function of the gap that is co-referential to the antecedent from which it is separated by an intervening RC or more. The use of an RP would facilitate such a task.

When we compare the Portuguese and Spanish multivariate analysis results for RPs, more differences arise. Factors that were significant in promoting RP occurrence in Portuguese, such as existential main clauses and RCs embedded within other subordinate clauses, did not affect the occurrence of RPs in the present study's data. Additionally, among the full prepositions that favor RPs in Spanish we have *para* (for, to) and *por* (for, to), whose Portuguese counterparts, in contrast, favor PP-chopping.

Corrected mean			0.048
Log likelihood			30.243
<hr/>			
Total N			152
	Factor Weight	%	N
<b>Animacy of the antecedent</b>			
Animate	.97	46.2	39
Inanimate	.23	2.7	113
<i>Range</i>	74		
<b>Semantic content of the preposition</b>			
Full	.96	42.9	14
Empty	.42	10.9	138
<i>Range</i>	54		
<b>Relative clause sequence</b>			
Stacked	.91	60	5
Single	.48	12.2	147
<i>Range</i>	43		
<b>Definiteness of the antecedent</b>			
Indefinite	.72	18.6	59
Definite	.36	10.8	93
<i>Range</i>	36		

<b>Position of the material</b>			
To the Left of the Verb	[.75]	22.2	27
At Both Sides of the Verb	[.48]	14.8	61
To the Right of the Verb	[.44]	11.1	45
No material	[.34]	5.3	19
<b>Restrictiveness of the relative clause</b>			
Non-restrictive	[.67]	20.9	43
Restrictive	[.43]	11	109
<b>Main clause semantics</b>			
Existential	[.62]	33.3	18
Other	[.48]	11.2	134
<b>Prepositional phrase status</b>			
Modifier	[.52]	12.3	122
Complement	[.42]	20	30
<b>RC position</b>			
Center	[.54]	14	43
Right	[.48]	13.8	109
<b>Number</b>			
Singular	[.52]	13.5	111
Plural	[.46]	14.6	41
<b>Clause embedding</b>			
Embedded	[.52]	4.5	22
Not embedded	[.50]	15.4	130

**Table 3.** Variable Rule Analysis of factors contributing to the occurrence of resumptive pronouns in oblique relative clauses in the Spanish of Lima, Peru

Table 4 shows the results of the factors favoring PP-chopping in Spanish. Contrary to my predictions, only a reduced set of factors, related to properties of the antecedent heads, was selected as significant: inanimate antecedents and singular antecedents favor the deletion of the PP. Moreover, factors that contributed significantly to the occurrence of PP-chopping in Portuguese, such as restrictive RCs and definite antecedents, along with all of the other factor groups, showed no effect on Spanish oblique RCs.

Corrected mean			0.327
Log likelihood			-92.354
Total N			152
	Factor Weight	%	N
<b>Animacy of the antecedent</b>			
Inanimate	.57	39.8	113
Animate	.30	17.9	39
<i>Range</i>	<i>27</i>		
<b>Number</b>			
Singular	.56	38.7	111
Plural	.35	22	41
<i>Range</i>	<i>21</i>		
<b>Relative clause sequence</b>			
Stacked	[.85]	40	5
Single	[.49]	34	147
<b>Position of the material</b>			
To the Right of the Verb	[.59]	40	45
To the Left of the Verb	[.56]	40.7	27
At Both Sides of the Verb	[.50]	32.8	61
No material	[.23]	15.8	19
<b>Restrictiveness of the relative clause</b>			
Restrictive	[.57]	36.7	109
Non-restrictive	[.34]	27.9	43
<b>Clause embedding</b>			
Embedded	[.71]	50	22
Not embedded	[.46]	31.5	130
<b>Prepositional phrase status</b>			
Complement	[.66]	43.3	30
Modifier	[.46]	32	122
<b>Main clause semantics</b>			
Other	[.52]	35.8	134
Existential	[.36]	22.2	18
<b>Semantic content of the preposition</b>			
Empty	[.51]	35.5	138
Full	[.38]	21.4	14

**Definiteness of the antecedent**

Indefinite	[.54]	35.6	59
Definite	[.47]	33.3	93
<b>RC position</b>			
Right	[.52]	34.9	109
Center	[.46]	32.6	43

**Table 4.** Variable Rule Analysis of factors contributing to the occurrence of prepositional phrase chopping in oblique relative clauses in the Spanish of Lima, Peru

It is worthwhile pointing out that inanimate antecedents entail the greatest probability of PP-chopping in an oblique RC. This is quite the opposite of the results seen in Table 3, where an animate antecedent is the factor with the strongest probabilistic weight in the production of RPs. I will return to the role of animacy in this linguistic variable later.

As for singular antecedent NP heads favoring PP-chopping, a feasible explanation would be that plural is the marked number, whereas singular is unmarked. This default condition of singular NPs may cause the PP that is co-referential with them to be more prone to being chopped than plural-marked NPs.

Since no factors related to the complexity in processing the function of the antecedent NP within the RC proved to be significant, in particular the case of 'light' RCs, it can be assumed that speakers do not resort to the PP-chopping strategy in RCs whose recoverability is easier in processing terms.

Finally, Table 5 presents the factors contributing to the occurrence of the standard variant, gap-containing RCs. There were only ten factors analyzed for this variant, since there were no instances of gaps in stacked RCs. As we can see in Table 5, 'light' RCs with no material inside other than the verb are the strongest predictor of gaps, a result that does not support the hypothesis based on the ease or difficulty of processing posed by the content of the RC: these clauses, which should present the simplest processing task, would be expected to have their PPs chopped more frequently, rather than favoring gap-containing structures. The second strongest factor contributing to the occurrence of gaps is an inanimate antecedent, which is a factor that also promotes PP-chopping.

Lastly, antecedent NPs whose heads are plural nouns are the third and weakest factor contributing to the production of a gap-containing oblique RC. Plural nouns are marked and they do not entail a default grammatical

number like singular ones do. As a consequence, NP antecedents with plural heads are not likely to have their co-referential PPs deleted, but rather maintained.

In general, the results of these three multivariate analyses show a completely different pattern when compared to the Brazilian Portuguese case. First, as mentioned before, the distribution of the variants is not similar: the structural preference in Spanish is gap > PP-chopping > RP, whereas in Brazilian Portuguese it is PP-chopping > RP > gap. Secondly, the constraints that govern the occurrence of RP and PP-chopping in Spanish are not in complementary distribution, and they do not coincide for the most part with the factors favoring their Portuguese counterparts.

Moreover, the loose connection that Tarallo (1986) based his explanation on does not seem to play an essential role in the choice of the three Spanish oblique RC variants. Instead, it appears that properties of the antecedent NP head, in particular, its animacy, are of crucial relevance in selecting one variant or another. As Tables 3-5 have shown, the strongest predictors of RCs containing RPs and PP-chopping are animate and inanimate antecedents, respectively, and the second strongest predictor of gap-containing RCs are also inanimate antecedents.

Corrected mean			0.526
Log likelihood			96.461
Total N			152
	Factor Weight	%	N
<b>Position of the material</b>			
No material	.80	78.9	19
At Both Sides of the Verb	.48	52.5	61
To the Right of the Verb	.47	48.9	45
To the Left of the Verb	.36	37	27
<i>Range</i>	44		
<b>Animacy of the antecedent</b>			
Inanimate	.56	57.5	113
Animate	.33	35.9	39
<i>Range</i>	23		
<b>Number</b>			
Plural	.64	63.4	41

Singular	.45	47.7	111
<i>Range</i>	19		
<b>Semantic content of the preposition</b>			
Empty	[.52]	53.6	138
Full	[.30]	35.7	14
<b>Prepositional phrase status</b>			
Modifier	[.54]	55.7	122
Complement	[.36]	36.7	30
<b>Clause embedding</b>			
Not embedded	[.52]	53.1	130
Embedded	[.37]	45.5	22
<b>Definiteness of the antecedent</b>			
Definite	[.55]	55.9	93
Indefinite	[.43]	45.8	59
<b>Restrictiveness of the relative clause</b>			
Non-restrictive	[.54]	51.2	43
Restrictive	[.49]	52.3	109
<b>RC position</b>			
Center	[.51]	53.5	43
Right	[.50]	51.4	109
<b>Main clause semantics</b>			
Existential	[.54]	44.4	18
Other	[.50]	53	134

**Table 5.** Variable Rule Analysis of factors contributing to the occurrence of gaps in oblique relative clauses in the Spanish of Lima, Peru

This twofold correlation of animate nouns and richer morpho-syntactic features on the one hand, and inanimate nouns and simpler morpho-syntactic marks on the other hand, goes along the lines of cross-linguistic patterns embodied in NP hierarchical orderings like Comrie's (1981) Animacy Hierarchy. In fact, it is acknowledged that in languages with a declension nominal system like Latin, Classical Greek, and Sanskrit, masculine and feminine nouns (generally referring to animate entities) show more abundant morphology than neuter nouns (generally referring to inanimate entities).

By the same token, Ancient Greek, Persian and Georgian use singular verbs with plural subject NPs with low animacy, and a number of languages have verbs agreeing with NPs higher in animacy rather than with NPs with



low animacy, sometimes overriding syntactic relations (Comrie 1981: 191). Furthermore, in some languages like Russian, masculine animate NPs have an overt accusative case marking in contrast with their inanimate counterparts, whose DO inflection coincides with the zero-marked nominative (Croft 2003: 131), and a similar pattern can be found in Polish (Yamamoto 1999: 49). In like manner, in Spanish, specific DO NPs with animate referents are marked by the preposition *a* (to), whereas nouns with inanimate referents lack such preposition<sup>13</sup> (Yamamoto 1999: 47).

On the one hand, the greater occurrence of RPs with animate antecedents<sup>14</sup> than with inanimate ones fits this cross-linguistic tendency; on the other, so does the correlation of inanimate antecedents and PP-chopping/gaps.

## 6. CONCLUSIONS

One could state, echoing Aristotle (2000 [350 B.C.]), that the golden mean or appropriate middle between two extremes –deficiency and excess– is desirable. In the Spanish oblique RCs case, this mean can be represented by the standard, gap-containing RCs, which in fact are the prevalent variant in the Peruvian variety analyzed, with 51.9% of occurrences. Deficiency, symbolized by PP-chopping, is the second most extended strategy, but at a rate of 34.2%, it is far from outscoring the gap-containing strategy. Excess, embodied in RP-containing RCs, is the least extended strategy, with only 13.8% of incidences of all the RCs analyzed.

These results suggest that Peruvian Spanish is conservative compared to Brazilian Portuguese with respect to the variable studied. The former has a considerable preference for the standard, gap-containing structure, whereas the latter shows a widely extended use of PP-chopping, a moderate occurrence of RPs, and a reduced presence of gap-containing RCs. In addition, as has been pointed out above, there is no complementary distribution between the non-standard variants studied in Spanish oblique RCs, in contrast to what Tarallo (1986) suggests occurs in Brazilian Portuguese.

<sup>13</sup> Non-specific animate DO NPs as ii) are not marked by the personal *a*, and the presence of the preposition entails a contrast in meaning:

- i) Busco *a* una secretaria bilingüe. 'I am looking for a bilingual secretary' (a specific one)
- ii) Busco una secretaria bilingüe. 'I am looking for a bilingual secretary' (any bilingual secretary)

<sup>14</sup> Recall that in Silva-Corvalán's (1996) study reviewed in Section 2, proper names, highly ranked in the Animacy Hierarchy, were the strongest factor conditioning the occurrence of RPs.

Regarding the first variant analyzed in this study, RPs, they are mainly favored by animate antecedents, in accordance with the richer morphology/agreement marks found cross-linguistically in NPs higher in the Animacy Hierarchy. The second strongest RP constraint is that of full prepositions, which I propose seem to stress their rich semantic content by promoting the use of an RP. Stacked RCs are the third factor favoring RPs, which speakers seem to use to more easily keep track of the function of an antecedent's trace in the RC, which might be obscured by one or more RCs intervening between them. Indefinite antecedents also favor the occurrence of RPs. The most plausible explanation for the effect of this constraint is that these antecedent NPs have an undetermined status in the main clause; however, when the speakers want to convey a determined status inside the RC to such an antecedent, they resort to using RPs as determiners where the gap co-indexed with the antecedent would appear, in a similar fashion in which definite articles are used (Trujillo 1990).

In general, compared to the results obtained in previous RP studies, we can conclude that the constraints favoring oblique RP cannot be drawn directly from the ones conditioning general<sup>15</sup> RP occurrence: the former are motivated by considerably less factors, among which is a constraint exclusive to oblique RCs, semantically full prepositions.

PP-chopping, on the other hand, is favored by inanimate antecedents, in accordance with cross-linguistic phenomena controlled by animacy, *i.e.* less morphology/syntactic relations for inanimate NPs. In addition, singular NPs, which represent the unmarked grammatical number and therefore, the least cognitively salient, are also the ones that are most frequently chopped in Spanish.

The third variant, the gap strategy, is mainly favored by RCs with no material within them except for the verb, *i.e.* light ones. However, this does not seem to imply that these light RCs pose the least effort in processing terms, since we would have expected them to favor PP-chopping, if they were in fact the easiest to process. Gaps are also favored by inanimate antecedents, following the patterns accounted for in the Animacy Hierarchy. In general, the lighter or heavier an RC is does not seem to ease or obscure, respectively, the processing of oblique RCs. The only processing difficulty that was found for this type of RCs was the stacking of clauses in a sequence, *i.e.* material outside the RC, not inside it.

---

<sup>15</sup> By *general*, I mean RPs as a whole, including subject and direct object RCs.

To sum up, the main contribution of this study is to have accounted for this Spanish morpho-syntactic variable comprising three variants, one of which had been only partially accounted for in previous studies. This analysis suggests that the factors conditioning oblique RPs are not the same ones that favor RPs performing different syntactic functions. Also, my findings indicate that the antecedent NP's animacy seems to play a critical role in choosing the variants of this three-way alternation, rather than the simplicity or complexity of the RC's argument structure.

## REFERENCES

- Alcina, Juan and José Manuel Blecua (1991): *Gramática española*. Barcelona: Ariel.
- Aristotle (2000 [350 B.C.]): *Nicomachean Ethics*. Cambridge: Cambridge University Press.  
<http://dx.doi.org/10.1017/CBO9780511802058>
- Bentivoglio, Paola (2003): “Las construcciones ‘de retoma’ en las cláusulas relativas: un análisis variacionista”, in Francisco Moreno, José Antonio Samper Padilla, María Vaquero, María Luz Gutiérrez Araus, César Hernández Alonso and Francisco Gimeno Menéndez (eds.): *Lengua, Variación y Contexto. Estudios dedicados a Humberto López Morales*, 2. Madrid: Arco Libros, pp. 507-520.
- Caravedo, Rocío (1989): *El español de Lima. Materiales para el estudio del habla culta*. Lima: Pontificia Universidad Católica del Perú, Fondo Editorial.
- Cerrón-Palomino, Álvaro (2006): “Pronombres de retoma en cláusulas relativas del castellano peruano: un fenómeno de causación múltiple”. *Lexis* 24 (2), pp. 231-258.
- (2011): “Simplicidad y claridad estructurales: la producción de cláusulas relativas en castellano”. *Lexis* 35 (2), pp. 225-259.
- Comrie, Bernard (1981): *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Croft, William (1996): “‘Markedness’ and ‘universals’: from the Prague school to typology”, in Kurt R. Jankowsky (ed.): *Multiple perspectives on the historical dimensions of language*. Münster: Nodus, pp. 15-21.
- (2003): *Typology and Universals*. Cambridge: Cambridge University Press.

- Herrera Santana, Juana (1990): "Duplicaciones pronominales en las oraciones de relativo", in María Ángeles Álvarez Martínez (ed.): *Actas del Congreso de la Sociedad Española de Lingüística. XX Aniversario*, 1. Madrid: Gredos, pp. 548-554.
- Samper, José, Clara Hernández and Otilia Pérez (2005): "Las construcciones de retoma en las cláusulas relativas en el español de Las Palmas de Gran Canaria", in César Hernández Alonso (ed.): *Filología y Lingüística. Estudios ofrecidos a Antonio Quilis*. Madrid: Consejo Superior de Investigaciones Científicas, Universidad Nacional de Educación a Distancia y Universidad de Valladolid, pp. 611-628.
- Silva-Corvalán, Carmen (1996): "Resumptive pronouns: a discourse explanation", in Claudia Parodi, Carlos Quicoli, Mario Saltarelli and María Luisa Zubizarreta (eds.): *Aspects of Romance Linguistics*. Washington D.C.: Georgetown University Press, pp. 383-395.
- Tarallo, Fernando (1986): "Functional and Structural Properties in a Variable Syntax", in David Sankoff (ed.): *Diversity and Diachrony*. Ottawa: John Benjamins, pp. 239-248.  
<http://dx.doi.org/10.1075/cilt.53.22tar>
- Trujillo, Ramón (1990): "Sobre la supuesta despronominalización del relativo". *Estudios de Lingüística* 6, pp. 23-46.
- Wurzel, Wolfgang (1988): "On Markedness". *Theoretical Linguistics* 24 (1), pp. 53-71.
- Yamamoto, Mutsumi (1999): *Animacy and Reference: A cognitive approach to corpus linguistics*. Amsterdam/Philadelphia: John Benjamins.  
<http://dx.doi.org/10.1075/slcs.46>

## Rare Phenomena and Representative Corpora

Bernard Comrie

*Max Planck Institute for Evolutionary Anthropology  
and University of California, Santa Barbara*

**Abstract:** Ditransitive constructions with a human theme, of the type *the teacher introduced his wife to the students*, constitute an interesting phenomenon in Spanish because of the behavior of “personal *a*”, with often conflicting elicited judgments and recommendations in grammars. An earlier corpus study carried out by the author, noting examples that occurred while reading Spanish-language novels, suggests that inclusion of personal *a* in such constructions is the usual, but not invariable, solution. However, occurrence of the construction is rare and uneven in texts, and it is difficult to search for it automatically in corpora. The reasons for this difficulty are set out, in particular the problem of identifying what string(s) should be searched for, as well as the advances, in particular in tagging, that would be needed to overcome them.

**Keywords:** Ditransitive constructions, Spanish syntax, corpus linguistics, linguistic methodology.



## 1. INTRODUCTION

It is not often that I abandon the shores of cross-linguistic typological investigations to venture into the glades of Ibero-Romance linguistics, but on occasion it happens, and on such occasions Carmen Silva-Corvalán's work is always there to serve as a guide. Perhaps my first major excursion into Ibero-Romance linguistics was Comrie (1982) on the syntax of clitic pronouns in written Brazilian Portuguese, for which I collected examples while reading Jorge Amado's *Terras do Sem Fim* (1942) ('The Violent Land' in its published English translation). At the time Carmen and I were colleagues at the University of Southern California, and she gave me invaluable comments on the manuscript, including saving me from some embarrassing faux-pas: I don't think I will ever again confuse "register" with "style".

In the intervening decades there have been many advances in the area of corpus-based linguistic research, and Carmen has been at the forefront in developing and applying such methods, whether making use of existing large corpora, such as the ARTHUS corpus in Eddington and Silva-Corvalán (2011) or creating her own sizable corpus for her research on bilingual and trilingual child language acquisition involving Spanish, as in Silva-Corvalán (2008).

In Comrie (2013) I returned to Spanish linguistics, to a problem that interested me both as a learner of Spanish and as a typologist, namely the use or non-use of "personal *a*" for a specific, human theme (T) argument in ditransitive constructions. I summarize this article in section 2, but mention here the method I used. Essentially, I continued with the method I had used for the 1982 paper, namely I noted down relevant examples that I encountered in my pleasure reading. Had I learned nothing in the meantime? Had developments in corpus linguistics entirely passed me by? I believe that the answer is "No", but justification will have to wait until section 3.

## 2. HUMAN THEMES IN SPANISH DITRANSITIVE CONSTRUCTIONS

Many textbooks and reference grammars of Spanish for foreigners warn learners that while the use of personal *a* is obligatory with specific, human patient (P) arguments of monotransitive verbs, as in *Mario no conoce a Elena* 'Mario does not know Elena', it should be avoided when there is a specific, human theme (T) of a ditransitive construction in the presence of a recipient



(R) also marked with the preposition *a*. This would disallow sentence (1a), while allowing (1b). (I use the symbol “#” to indicate that the grammaticality or acceptability of the example in question is at issue, without taking a stand on whether or not the example is grammatical or acceptable. In a ditransitive construction such as *le envió el dinero a su hija* ‘he sent the money to his daughter’, the constituent in the position of *el dinero* is the theme, that in the position of *a su hija* the recipient.)

- (1a) #El maestro (les) presentó a su mujer a los alumnos.  
 (1b) #El maestro (les) presentó su mujer a los alumnos.  
 ‘The teacher introduced his wife to the students’.

Before embarking on the research that led to my 2013 article, I had noticed in passing what struck me as exceptions to the advice given to foreign learners. However, my attempts to gain a clear understanding of the situation by asking native speakers did not produce any clear result, with reactions ranging from the type “of course examples like (1a) are fine – how else would you express it?” to those of the type “No, I would definitely avoid examples like (1a)”. It was time to examine actual usage. Given my resources at the time, this meant essentially examining written usage, as indicated in section 1 by noting examples from my reading. As described in more detail in Comrie (2013), reading about 2,000 pages of Spanish yielded a scant five clear examples, four of type (1a) and one of type (1b). Rather than repeating material from the earlier article, I cite two relevant examples from my more recent reading of *El sueño del celta* (2010) (‘The Dream of the Celt’, which is also the title of the official English translation) by the Peruvian writer Mario Vargas Llosa. In (2), the theme takes personal *a*, while in (3) it does not.

- (2) El prefecto fue presentando a Roger [T] a los demás invitados [P]: ...  
 ‘The prefect introduced Roger to the other guests: ...’  
 (3) Encargó éstos [= los barbadenses] a Frederick Bishop, ...  
 ‘He entrusted the latter [= the Barbadians] to Frederick Bishop, ...’

My conclusion in the 2013 article, based also on consideration of examples with pronominal recipients, was that the expression of specific, human themes by means of personal *a* is dominant in the usage of at least the sources I examined, as in (1a), but that omission of personal *a*, as in (1b), is a less preferred alternative, even though that noun phrase would

have required personal *a* as a patient of a monotransitive verb. In the 2013 article I restricted myself more or less to the empirical conclusion and its relevance for the typology of ditransitive constructions, although I did allude to possible conclusions relating to methodology, and it is to these that I now turn.

### 3. METHODOLOGICAL IMPLICATIONS

The first point to note with regard to the incidence of ditransitive constructions with specific, human themes is that they are rare – perhaps twenty-two clear examples in about 2,000 printed pages. (The hedge “perhaps” is needed because I excluded a few examples on the grounds that their themes were, though human, probably non-specific, although it is not always easy to be confident in such interpretative judgments.) Now, rarity in itself is not a problem for corpus linguistics, since if one has a large enough corpus, then even rare phenomena that have a more or less even distribution throughout texts will appear in sufficient quantity to be amenable to standard corpus linguistic methodology. The problem with animate themes is that they are not evenly distributed across texts, not even remotely so. Texts or portions of texts dealing with certain topics are much more likely to give rise to such examples. Indeed, of the books that form the empirical basis of the 2013 article, by far the greatest incidence of relevant examples comes from *La catedral del mar* (2006) (‘The Cathedral of the Sea’, also the title of the official English translation) by Ildefonso Falcones, a Catalan author writing in Spanish. The novel deals with medieval Catalonia, and its locale gives rise to far more instances of giving people to other people than one would expect in a text dealing with modern society, in addition to which the subject matter also includes specific passages that increase the frequency of the relevant construction, such as giving babies to adults (to hold), giving brides in marriage, handing people over to judicial authorities. It is thus not simply the case that ditransitive constructions with human themes are rare, rather the occasions on which it would be relevant to use them (at least outside a linguistics article) are rare.

With lexical items, rarity and uneven distribution in a representative corpus are not necessarily unexpected. If one takes a lexical item like *iliaco* (variant: *ilíaco*) or its English equivalent *iliac*, an adjective derived from the name of one of the bones constituting the pelvis, then one would expect

precisely such a distribution with a technical medical term, and would probably reduce the corpus to its medical sub-corpus in order to investigate the term further. But one does not expect this with grammatical, whether morphological or syntactic, phenomena, unless of course they are tied to particular lexical items, such as the idiosyncratic morpho-lexical stress changes found in the items *régimen* ‘regime’ (plural *regímenes*) and *carácter* ‘character’ (plural *caracteres*). However, in languages with a large number of ditransitive verbs like Spanish and English, the phenomenon of ditransitive verbs with human themes is not tied to any particular lexical item, yet nonetheless the incidence is low and uneven. One might of course expect to find a higher incidence in texts dealing with certain topics, such as marriage negotiations or slave-trading, but restricting one’s attention to such texts would equally exclude a portion of the examples that are attested, with the possibility that results might be specific to descriptions of these specific situations and not generalizable to all uses of ditransitive verbs with human themes.

Of course, with an example like *iliaco/ilíaco* it would be possible to search the whole of a large corpus, because it is obvious what string one is searching for, even if the computing effort is great in relation to the number of examples eventually found. With ditransitive constructions with human themes, however, it is unclear what string one would search for.<sup>1</sup> In principle one could extract from a corpus all sentences containing a ditransitive verb, and this would mean fewer sentences to read than my technique of reading novel after novel, but I suspect that reading through all the disjointed ditransitive clauses extracted from a corpus would be considerably less interesting than reading a novel. One might suspect that at least some sequences might be retrievable automatically, but even here there are problems. The sequence *presentó su mujer* from (1b), repeated for various ditransitive verbs, might seem a plausible way of extracting some relevant examples, but given the frequency of the order Verb – Subject in Spanish, it would be far from foolproof. An example with verb and subject distinct in number, such as *presentó sus hijos* ‘he introduced his sons’ might work in getting only relevant examples, but would of course fall far short of getting all relevant examples. A string search seems to have little hope of improvement over the novel-reading approach.

---

<sup>1</sup> I am grateful to John Du Bois and Dan Slobin for discussion of this point, a discussion which reinforced my pessimism, although I of course bear full responsibility for this interpretation.

The last possibility would be a tagged corpus. Let us be clear on what would need to be tagged. Noun phrases and prepositional phrases with the preposition *a* would need to be tagged according to whether they constitute a theme or not, and the latter would further need to be tagged as to whether they constitute a recipient or not. (It would also be necessary to tag head nouns as human or not, a task that should be readily feasible in tandem with a dictionary that identifies lexical items as having human or non-human reference.) If done manually, this would be as time-consuming as reading through the text and extracting the examples as I did in the research summarized in section 2. The only hope that I see for automating the procedure is to implement an automatic parser that would, with some high degree of reliability, identify themes and recipients in Spanish text. Is this feasible?

#### 4. CONCLUSIONS

On the one hand, I hope to have shown that I am still capable of learning when it comes to the contributions of corpus linguistics – *lernfähig* ‘learn-capable’, as my German-speaking colleagues would succinctly say, which is not quite the same as *teachable*, since one can learn without being taught. I recognize more than ever that corpus linguistics can play an important role in the kinds of typological investigations that form the center of my research activities. But in the case of human themes of ditransitive verbs, I feel that I have gone as far as I can in using the methodology, which amounts essentially to not developing it any further than what was available at the dawn of literacy: Read and note relevant examples. But I look forward to developments in corpus linguistics and related disciplines that may overcome the limitations that beset my investigation. Any suggestions, Carmen?



## REFERENCES

- Comrie, Bernard (1982): “Remarks on clitic-climbing in Brazilian Portuguese”, *Lingua* 58, pp. 243-265.
- (2013): “Human themes in Spanish ditransitive constructions”, in Dik Bakker and Martin Haspelmath (eds.): *Languages Across Boundaries: Studies in Memory of Anna Siewierska*. Berlin: De Gruyter Mouton, pp. 37-52.
- Eddington, David; Silva-Corvalán, Carmen (2011): “Variation in the use of *deber* and *deber de* in written and oral materials from Latin America and Spain”, *Spanish in Contact* 8 (2), pp. 257-271.  
<http://dx.doi.org/10.1075/sic.8.2.04edd>
- Silva-Corvalán, Carmen (2008): “La adquisición bilingüe (y trilingüe): aspectos sintácticos, morfosintácticos y léxicos”, in Rosa María Ortiz Ciscomani (ed.): *IX Encuentro Internacional de Lingüística en el Noroeste*. Hermosillo, Sonora: Editorial Unison, pp. 13-34.





## **Subject Position in Spanish: a Study of Factor Interactions with Prototypical Verbs**

Roberto Mayoral Hernández  
*University of Alabama at Birmingham*

**Abstract:** This paper is primarily concerned with subject position in Spanish. While there is abundant research that deals with subject expression in Spanish, only recently there has been an increased interest in subject position. This article uses the variationist methodology to study hypotheses that relate to the language faculty, such as the Unaccusative Hypothesis (Perlmutter 1978) and the role of animacy in argument structure (Comrie 1989). The variables studied are: *verb type* (transitive, unaccusative, unergative), *type of clause* (main, completive, relative), *presence of co-occurring postverbal XP* (preverbal, postverbal), *presence of co-occurring preverbal XP* (preverbal, postverbal), and *subject animacy* (animate, inanimate). The program Goldvarb X was utilized to run regression analyses and the Pearson's chi-square test was used to study relationships between factors. The results indicate that all the independent variables have significant effects on subject position except for animacy. Furthermore, the subjects of transitive verbs were shown to be predominantly inanimate.

**Keywords:** unaccusative, animacy, verbs, variation, subject.





## 1. INTRODUCTION

The study of the variable position of subjects in Spanish is receiving an increasing attention within the field of Hispanic linguistics. There is abundant research regarding the topic of subject expression, which has been carried out from different perspectives and disciplines, such as sociolinguistics (Silva-Corvalán 2003; Poplack 1980), pragmatics (Prince 1981), syntax (Torrego 1989) and dialectology (Lipski 1977). However, the approaches used to address this topic vary greatly. While syntactocentric studies (Torrego 1989 *inter alia*) are concerned with underlying structures of Universal Grammar and rely on native intuitions to gather data, variationist studies focus on specific language varieties (Díaz Campos 2001, Bel 2003), and use statistical methods and field studies to compile and analyze data. Linguistic frameworks with focus on syntax, such as the Generative Grammar and Relational Grammar, are concerned with the language faculty, while the studies of variation study performance. Even though the topics that are discussed in syntactocentric frameworks are sometimes very different from those addressed using variationist approaches, it is possible to find a common ground and utilize statistical methods to further corroborate linguistic hypotheses. This paper investigates syntactic theories, such as the validity of the Unaccusative Hypothesis (Perlmutter 1978), using the variationist approach to describe Spanish verbs and the phenomenon of subject inversion; a semantic factor, such as the role of animacy in subject expression (Comrie 1989); and the collocation of constituents in the sentence, which is a factor that have been studied from different disciplines, such as syntax, psycholinguistics and pragmatics. The methodology adopted is the variationist approach, because it is well suited to analyzing variable contexts and quantifying the interaction of different variables. My belief is that the different linguistic disciplines can benefit from each other and contribute to the discovery of new principles and the advancement of the linguistic field.

This paper presents a corpus analysis of subject position in Spanish with data extracted from the online *Corpus de Referencia del Español Actual* (Reference Corpus of Contemporary Spanish). This Spanish corpus, which is normally named using its acronym CREA, is the most extensive corpus of Spanish. The biggest part of CREA is formed by written data, although other oral corpora have been more recently added to the database. Unlike other studies that deal with subject position using the variationist approach, this paper deals with written Spanish. Several factors that may influence subject position in Spanish, which may be preverbal or postverbal, are analyzed

here: (i) *verb type* (transitive, unaccusative, unergative), (ii) *presence of co-occurring postverbal XP* (absence of XP, presence of XP), (iii) *presence of co-occurring preverbal XP* (absence of XP, presence of XP), (iv) *type of clause* (main, completive, relative), and (v) *subject animacy* (animate, inanimate). Variables (i) and (iv) are syntactic in nature. Variables (ii) and (iii) refer to the distribution of non-subject constituents in the sentence, and could be described as syntactic, psycholinguistic or pragmatic phenomena, depending on the framework adopted. Finally, subject animacy is a lexical-semantic factor which has been claimed to influence how arguments will surface into syntax (Comrie 1989).

Several statistical tests have been utilized in order to determine whether factors are significant or not. Multiple cross tabulations analyzed with the Pearson's chi-square test were carried out to study the distribution of subject position and how it is influenced by the five independent variables introduced above. The program Goldvarb X was utilized to run regression analyses and determine the best fitting model.

The structure of the paper is as follows. In section 2, the properties of Spanish subjects are briefly described. Section 3 introduces several factors on which the literature dealing with subject position has focused. Section 4 develops the corpus analysis carried out here and presents the results. Section 5 includes a discussion of the results, and section 6 provides a brief conclusion.

## 2. PROPERTIES OF SPANISH SUBJECTS

The term subject has been used in the linguistics literature to describe different grammatical entities. In this paper, subjects are described as the argument that overtly agrees with the verb. Subjects are the only arguments in Spanish that express overt gender and number agreement with the verb. Since Spanish is a pro-drop language, subjects can be omitted under certain conditions, although the contexts where omission can take place are not random (Silva-Corvalán 1982). The environments where subject pronouns can be omitted are not only subject to general grammatical constraints, but are also influenced by other factors, such as the variety of Spanish or dialect being analyzed (Orozco and Guy 2008)<sup>1</sup>.

<sup>1</sup> For example, it is a well known fact that many varieties of Caribbean Spanish have a higher percentage of subject expression in the second person singular, but only in certain constructions, creating thus sentences that would not occur in other varieties of Spanish.

As far as subject position is concerned, the unmarked position is preverbal, but subjects can also occur in postverbal position with both transitive and intransitive verbs in declarative sentences. The sentences in (1) include the transitive verb *empujar* “push”, while (2) and (3) are examples of intransitive verbs. In particular, the verb used in (2) is the unergative verb *trabajar* “work”, and the one in (3) is the unaccusative *llegar* “arrive”.

- (1) a. Un niño chino empuja su carrito en la ciudad de Kunming.  
A child Chinese pushes his cart in the city of Kunming  
“A Chinese child pushes his cart in the city of Kunming.”  
(Press, “China anuncia una ley para “evitar los nacimientos de calidad inferior””, *El Mundo* (January 10, 1994); Spain)
- b. Sólo empuja el libro desde el borde de la mesa, aquel que tensiona la espalda  
Only pushes the book from the edge of the table, that who bends the back  
“Only the one who is bending his back is pushing the book from the table end”  
(J.M. Romero Blanch, *El arte de la equitación* (1990: 48); Argentina)
- (2) a. El minero trabaja durante toda la noche y...  
The miner works during all the night and...  
“The miner works during the whole night and...”  
(Moema Viezzer, *Si me permiten hablar...* (1977: 26); Bolivia)
- b. En el coro trabaja un maestro de tracerías y taraceas.  
In the choir works a master of tracery and marquetry  
“A master of tracery and marquetry works in the choir”.  
(Victor Chamorro, *El muerto resucitado* (1984: 33); Spain)
- (3) a. En 1896, Picasso llega a Barcelona.  
In 1896, Picasso arrives to Barcelona  
“In 1896, Picasso arrived in Barcelona”.  
(Cesáreo Rodríguez-Aguilera, *Arte moderno en Cataluña* (1986); Spain)

- b. En febrero llega el grupo Queen a la Argentina.  
 in February arrives the group Queen to the Argentina  
 “In February, the band *Queen* will arrive in Argentina”.  
 (Laura Ramos and Cynthia Lejbowicz, *Corazones en llamas. Historias del rock argentino en los “80* (1992: 37); Argentina)

The previous examples show strong variation in subject position, since subjects may occur in preverbal and postverbal positions, either immediately adjacent to the verb or separated from it by one or more phrases. But what determines the position of subjects? Variationist studies have proposed different factors that have been shown to influence subject position. The following section describes some factors that are relevant for the analysis presented here.

### 3. FACTORS THAT INFLUENCE SUBJECT POSITION

In the previous section, subject position has been shown to be very variable in Spanish, as seen in examples (1), (2) and (3). While it is true that constituent ordering in Spanish is much more flexible than in languages like English, it is also true that position is constrained by different factors. One of the factors that have received a most detailed analysis is information structure, which is normally considered a dichotomous variable with two categories, which are typically called *topic-focus* or *theme-rheme* (see, e.g., Ocampo 2003, Zubizarreta 1998). These categories broadly correspond to the binary distinction between new and given information, where topic and theme correspond to old or given information while focus and rheme represent new information. However, the apparent simplicity of this factor is not real. Numerous studies have gone away from the new/given distinction and adopted a much more complicated division of the information structure factor, with many different degrees of givenness (Prince 1981). In spite of the different methodologies used to define and categorize information structure, the general observation is that new information tends to appear in postverbal position in Spanish, while given information tends to precede newer information or be preverbal. Silva-Corvalán (1982) shows that constituents in a sentence tend to be linearly ordered depending on their degree of givenness, so that older constituents will precede newer ones.

The study of information structure finds a procedural obstacle at the time of determining what is new or given information. Firbas (1964) proposes that givenness is not a dichotomous variable, but rather a scale or hierarchy, which he calls Communicative Dynamism. Building on this idea, Chafe (1976, 1987, 1994) reinterprets givenness as a cognitive or psychological category, and postulates that using the term *activation* would be more adequate than the former terms of given and new. As seen above, the binary distinction is no longer favored in later studies, and different givenness scales were developed, with different degrees of complexity, such as the Scale of Familiarity (Prince 1981), the Givenness Hierarchy (Gundel, Hedberg and Zacharski 1993) or the Accessibility Marking Scale (Ariel 1990). As the studies on givenness drift towards the assumption of underlying psycholinguistic and mental processes, such as processing, degrees of activation, and lexical retrieval, it becomes increasingly more difficult for the human annotator to accurately code a text (see for example Lambrecht 1994 and Kahn and Arnold 2012 for a more detailed discussion).

While being an extremely important variable, information structure by itself cannot account for the distribution of subjects in Spanish. For example, transitive verbs seem to have an almost categorical SV ordering, while intransitive verbs, whether they are unaccusative or unergative show a higher percentage of postverbal subjects than transitive verbs. The following sections introduce the different variables that are going to be examined in this paper.

### 3.1. Verb types

Because verbs play a central role in the sentence, they have been studied from many different perspectives. Mayoral Hernández 2007 showed that the distribution of subjects is significantly influenced by the type of verb. In particular, the author analyzed three main verb types: (i) transitive, (ii) unergative and (iii) unaccusative. Examples of sentences containing these three types are provided in (1), (2), and (3) respectively.

Using the framework of Relational Grammar, Perlmutter 1978 postulated the existence of two different types of intransitive verbs, which he called unergatives and unaccusatives, developing in that manner the Unaccusative Hypothesis. These two types of intransitive verbs differ in their semantics and their syntax, as they react differently to particular syntactic tests. From a semantic point of view, unergative verbs have subjects with agentive properties, while the subject of unaccusatives has more patient properties (Perlmutter 1978; Dowty 1991). As far as their syntax is concerned,

the subject of unaccusatives has been described as being an underlying object in those syntactic theories that provide different levels of interpretation, such as Relational Grammar and the Generative tradition. However, unergative verbs are claimed to have “real” subjects. These observations have been supported by abundant crosslinguistic syntactic evidence. For example, Burzio (1981, 1986) noted how unaccusatives and unergatives behave differently to *ne-cliticization* and only the subjects of unaccusatives display overt agreement with the past participle in complex tenses in Italian and French. Furthermore, Sorace (2000) compiled abundant information regarding the selection of auxiliaries by these two types of intransitive verbs in different languages, which led her to postulate the existence of an auxiliary selection hierarchy, where typically unergative verbs select mostly the auxiliary “have” in complex periphrastic forms, while typical unaccusatives select the auxiliary “be”.

A problem with this theory is the identification of the specific set of verbs that can be identified either as unergative or unaccusative in a given language (see Mendikoetxea 2000 for a classification of unaccusative verbs in Spanish). There are some clear cases that seem to have a stable crosslinguistic classification, but there are also verbs that can behave as either unaccusative or unergative depending on the language and also on the construction where they occur (for a discussion of some of these constructions with unaccusative properties see Zubizarreta and Oh 2007 and Mayoral Hernández 2008 and 2010).

### **3.2. Presence of co-occurring XPs**

The linguistic literature dealing with weight effects and information structure have analyzed how the presence of other constituents affects the collocation of co-occurring phrases (phrases *x* or XPs) under certain circumstances. For example, Hawkins (2004) observes how an increase in weight, i.e., an increase in the number of words, is associated with a higher chance for an XP to be placed in sentence final position. In his studies, he shows how the weight of the different constituents in a given sentence interact with each other to favor an ordering where heavier elements occur sentence finally, preceded by lighter elements. This is not a rule though, it is a preference.

Although Hawkins has only studied constituent ordering in English in postverbal positions, it is possible to hypothesize that constituent ordering in Spanish might also be affected by the presence of other XPs.

### 3.3. Type of clause

Another factor that has been claimed to influence subject position is the type of clause. Subject inversion, has been associated with relative clauses in different studies (see for example Gutiérrez-Bravo 2003 and Butt and Benjamin 2011). Mayoral Hernández (2008) quantified subject position with different types of sentences, and observed that relative clauses occur with postverbal subjects in 76% of the cases.

### 3.4. Animacy

The topic of animacy has received particular attention within the area of typology, and has been claimed to be partially responsible for the existence of ergative and nominative languages, including split ergativity in languages like Basque (Comrie 1989). Comrie proposed an animacy hierarchy that would determine, in conjunction with other factors, what will surface as subject or object in a given language: human > animate > inanimate. In this hierarchy, for example, a human NP will tend to surface as nominative in a nominative-accusative language, and therefore as a sentential subject, more easily than an inanimate NP. So what would happen with inanimate subjects? Will they appear in postverbal position because that is the default position for objects in Spanish? This is one of the questions that will be addressed in this paper.

Animacy has been defined in different ways. Beyond the categories of *human*, *animate* and *inanimate*, other subclassifications, such as “government”, have been postulated. Government has been proposed in order to identify institutions created by humans which have a collective interpretation.

## 4. ANALYSIS OF FACTORS

This section presents the experiment that was developed to analyze the interaction of the different variables, and discusses the results.

The dependent variable is subject position, coded as a dichotomous variable with two different levels: preverbal and postverbal. The study includes 5 independent variables: (i) *verb type*, with the values *transitive*, *unaccusative*, and *unergative*; (ii) *presence of co-occurring preverbal XP*, with the attributes *presence of XP* and *absence of XP*; (iii) *presence of co-occurring postverbal XP*, with the attributes *presence of XP* and *absence of XP*; and (iv) *type of clause*, with the values *main clause*, *relative clause*, *completive clause*; and (v) *subject animacy*, with the values *animate* and *inanimate*



The corpus sample used for this study is based on Mayoral Hernández (2007), and therefore the same coding method has been followed. It is formed by 450 sentences extracted from the online corpus *Corpus de Referencia del Español Actual* (CREA). This corpus CREA contains mostly sentences extracted from written sources, although there are some oral corpora that have been added to the main corpus. The corpus sample used for the current study contains sentences from written media which contain overt subjects. Out of all the sentences in the sample, only 5.8% have a pronominal subject. Only declarative sentences are included, since interrogatives imply processes of subject inversion that would compromise the reliability of the analysis.

Section 3.1 addresses one of the shortcomings of adopting a syntactic division of verb types, since it is sometimes unclear if some verbs should be described as either unergative or unaccusative. In order to avoid this problem, only verbs with an uncontroversial classification have been selected. In particular, *empujar* “push” represents transitive verbs, while *trabajar* “work” and *llegar* “arrive” represent unergatives and unaccusatives respectively. In order to avoid the interference of factors such as number, person or aspect, only sentences containing the third person singular of the present indicative, active voice have been selected. Another coding decision necessary to complete the annotation involves determining when a verb should be considered transitive. It is a well known fact that even typically transitive verbs can be used in an intransitive manner. Because of this, only those sentences containing an overt object were coded as transitive. Said object could be realized either as a clitic pronoun or as an NP.

When coding the co-occurrence of other XPs in the sentence, clitic pronouns were ignored, as well as any other particle with a non-variable position. Therefore, a sentence containing the structure *conjunction + subject + clitic + verb* would be coded as not having any co-occurring XPs.

The *type of clause* variable identifies the clause where the alternating verb appears. Sentences were coded as *main clause* when they were part of a sentence with no introductory nexus or complementizer, or when the sentence was introduced by a coordinating conjunction, such as *and* or *but*. Sentences were coded as relative clause when they were introduced by a relative pronoun or *wh-* word fulfilling a function in the sentence other than nexus, such as subject, object or adjunct. For example, the pronoun *that* in the sentence *I ate the ice-cream that<sub>i</sub> you prepared t<sub>i</sub>* functions as the direct object of *prepare*, so it would be coded as a relative clause. Finally,

embedded clauses were coded as completive when they were headed by a complementizer without a syntactic function, such as the conjunction *that* in the sentence *I said **that** I want an ice-cream.*

As far as subject animacy is concerned, this variable was initially coded using three possible values: *human*, *animate non-human*, and *inanimate*. However, there were only four non-human animate subjects in the whole corpus sample, so both human and non-human animate subjects were recoded into the broader category *animate*, becoming thus a binary independent variable.

The corpus sample has been analyzed using a series of crosstabs to select the right coding, applying the Pearson's chi-square statistics. Eventually, a binomial logistic regression was run utilizing the statistical software Goldvarb X in order to determine the best model and the relative weight of every factor.

#### 4.1. Results: Verb type

The crosstabulation in table 1 indicates that there exists a statistically significant relationship between subject position and verb type ( $p < 0.05$ ).

			Verb type			Total
			Transitive	Unergative	Unaccusative	
Subject Position	Postverbal	Count	5	32	62	99
		Column %	3.0%	21.6%	45.9%	22.0%
	Preverbal	Count	162	116	73	351
		Column %	97.0%	78.4%	54.1%	78.0%
Total		Count	167	148	135	450
		Column %	100%	100%	100%	100%
		Value	df	p value		
Pearson's Chi-Square		80.202	2	.0001		

Table 1: Subject position by Verb type

Both transitive and unergative verbs display a clear preference for preverbal subjects, although there is also a statistically significant difference between them. In fact, while the subjects of transitive verbs occur in an almost

categorical preverbal distribution, unergatives still present a significant amount of postverbal subjects. By contrast, the subjects of unaccusative verbs are distributed almost evenly between the preverbal and postverbal domains.

**4.2. Results: Presence of co-occurring preverbal XP**

The presence of other constituents in preverbal position has also a statistically significant influence on subject position in Spanish ( $p < 0.05$ ). The data displayed in table 2 indicate that the presence of a co-occurring preverbal XP increases the chance of having a postverbal subject to 34.4%. However, when there is no co-occurring preverbal XP, the percentage of postverbal subjects overall falls to only 17.4%.

		Presence of preverbal XP			Total
		No co-occurring preverbal XP	Co-occurring preverbal XP		
Subject Position	Postverbal	Count	57	42	99
		Column %	17.4%	34.4%	22.0%
	Preverbal	Count	271	80	351
		Column %	82.6%	65.6%	78.0%
Total		Count	328	122	450
		Column %	100%	100%	100%
		Value	df	p value	
Pearson's Chi-Square		15.061	1	.0001	

Table 2. Subject position by Presence of co-occurring preverbal XP

**4.3. Results: Presence of co-occurring postverbal XP**

The data observed in table 3 reveal that the presence of a constituent in postverbal position significantly increases the percentage of preverbal subjects. In fact, when a postverbal XP occurs, the subject will be preverbal in 92.7% of cases. However, if there is no other constituent following the verb, the subject will become postverbal in 69.2% of cases, while it will occur in the preverbal domain only in 30.8% of the total occurrences. The chi-square test indicates that this dependent variable is statistically significant ( $p < 0.05$ ).

			Presence of preverbal XP		Total
			No co-occurring preverbal XP	Co-occurring preverbal XP	
Subject Position	Postverbal	Count	74	25	99
		Column %	69.2%	7.3%	22.0%
	Preverbal	Count	33	318	351
		Column %	30.8%	92.7%	78.0%
Total		Count	107	343	450
		Column %	100%	100%	100%
		Value	df	p value	
Pearson's Chi-Square		181.933	1	.0001	

Table 3. Subject position by Presence of co-occurring postverbal XP

#### 4.4. Results: Type of clause

The type of clause is a statistically significant factor as well ( $p < 0.05$ ). However, the distribution of subject position in completive and main clauses is almost identical, with an obvious tendency to place subjects in the preverbal domain, and in fact there is no statistically significant difference between main and completive clauses as far as subject position is concerned ( $p = 0.717$ ).

			Type of clause			Total
			Completive	Main	Relative	
Subject Position	Postverbal	Count	7	41	51	99
		Column %	12.1%	13.9%	53.1%	22.0%
	Preverbal	Count	51	255	45	351
		Column %	87.9%	86.1%	46.9%	78.0%
Total		Count	58	296	96	450
		Column %	100%	100%	100%	100%
		Value	df	p value		
Pearson's Chi-Square		68.984	2	.0001		

Table 4. Subject position by Type of clause

Table 4 shows that the overall influence of the type of clause factor is indeed significant, because relative clauses, unlike the other two types, show no clear preference for either preverbal or postverbal subjects, with maybe a slight preference to place subjects in the postverbal domain (53.1%).

#### 4.5. Results: Subject animacy

The overall effect of subject animacy is not a significant factor ( $p > 0.05$ ). Table 5 shows that while inanimate subjects have a slightly higher percentage of postverbal occurrences (25.8%) when compared to animate entities (18.6%), such difference is not statistically significant.

			Subject animacy		Total
			Animate	Inanimate	
Subject Position	Postverbal	Count	44	55	99
		Column %	18.6%	25.8%	22.0%
	Preverbal	Count	193	158	351
		Column %	81.4%	74.2%	78.0%
Total		Count	237	213	450
		Column %	100%	100%	100%
		Value	df	p value	
Pearson's Chi-Square		3.442	1	.064	

Table 5. Subject position by Subject animacy

This is due to the fact that transitive and unaccusative verbs pattern together as far as subject animacy is concerned. Table 6 reveals that both transitive and unaccusative verbs have a very high and strikingly similar percentage of inanimate subjects, 62.3% and 63% respectively, while unergatives display the opposite tendency, with only 16.2%.

		Verb type			Total	
		Unergative	Transitive	Unaccusative		
Subject Animacy	Animate	Count	124	63	50	237
		Column %	83.8%	37.7%	37.0%	52.7%
	Inanimate	Count	24	104	85	213
		Column %	16.2%	62.3%	63.0%	47.3%
Total		Count	148	167	135	450
		Column %	100%	100%	100%	100%
		Value	df	p value		
Pearson's Chi-Square		85.671	2	.0001		

Table 6. Animacy by Verb type

Furthermore, the effect of subject animacy on subject position is only statistically significant with unaccusative verbs in the corpus sample.

Furthermore, when crosstabs designed to examine the interrelation between subject position and subject animacy are run for the three verb types individually, the effect of subject animacy turns statistically significant only with unaccusatives in the corpus sample.

#### 4.6. Factor weights

This section examines the weight of the different factor groups by carrying out a binomial logistic regression (binomial up and down) with the statistical software Goldvarb X.

Factor group	Factor weight	Factor group	Factor weight
Verb type		Presence of postverbal XP	
Unaccusative	.80	No co-occurring XP	.91
Unergative	.56	Co-occurring XP	.32
Transitive	.21	Range	59
Range	59	Clause type	
Presence of preverbal XP		Relative	.76
No co-occurring XP	.38	Main	.40
Co-occurring XP	.79	Completive	.56
Range	41	Range	36
Log likelihood	-117.461	p value	0.001

Table 7. Factor weights determining the occurrence of postverbal subjects

The regression analysis considered verb type, presence of postverbal XP, presence of preverbal XP and clause type as statistically significant, and added them to the best model (table 7). Table 7 indicates that verb type and the presence of a co-occurring postverbal XP are the most relevant factor groups, since they have the highest range (range = 59), followed by presence of preverbal XP (range = 41), and clause type (range = 36). In agreement with the results observed when using the  $\chi^2$  tests, Goldvarb excluded subject animacy from the model.

If transitive verbs are eliminated to compare only the two types of intransitive verbs, animacy is still excluded from the model during the regression.

Regarding the particular factors that cause the occurrence of postverbal subjects, the absence of a co-occurring postverbal XP is the most important (weight = .91), followed by the presence of an unaccusative verb (weight = .80), the presence of a preverbal XP (weight = .79) and relative clauses (weight = .76).

## 5. DISCUSSION

The data presented in section 4 show that the distribution of subject position is significantly influenced by several factors. The syntactic factor verb type is extremely important here, as it can be used to predict subject position very reliably, especially for transitive verbs, which show an almost categorical preference for preverbal position. Unaccusatives are the verb type with a higher percentage of postverbal occurrences. Similar observations have been made in different studies, as mentioned in section 3.1, and it is expected from a theoretical point of view if one considers that unaccusative verbs are underlying objects following the Unaccusative Hypothesis. This specific factor presents a point of convergence between syntactocentric frameworks and the variationist approach, since the latter has been used to further corroborate the existence of a difference between unergative and unaccusative verbs. There is, however, at least one paper where the unergative/unaccusative distinction has not been found to be significant: Ortiz López (2010). In his study, Ortiz López showed a series of question/answer conversations to monolingual and bilingual speakers, who were then asked to judge the grammaticality of the different sentences using a 1 to 5 scale. Therefore, this study was based on grammaticality or acceptability

judgments. Psycholinguistic research suggests that there are many problems related to the validity of acceptability judgments. While they are a tool that can be used to evaluate language competence, they may result in sentences which, while possible, will never be found in spontaneous speech or even more formal written language. Another possible intervening factor here is that the variety of Spanish analyzed by Oztiz López, Caribbean Spanish, is known to have a higher preference for preverbal subjects, even in contexts that would be ungrammatical in other varieties, such as direct questions. At this point, it would be interesting to see research on Caribbean Spanish that relies on language gathered from either oral or written texts, eliminating the flaws of acceptability judgments. If Ortiz López's interesting results were confirmed by further analyses, it could indicate that Caribbean Spanish might be evolving to a more syntactically simple system, more similar to French or even English.

The presence or absence of co-occurring XPs has been proven to be statistically significant, especially in postverbal positions. While the absence of postverbal XP increases the chances of having a postverbal subject, the absence of a preverbal XP increases the chance of subjects in the preverbal domain. This factor might be related to how the human processor organizes information as it becomes active in the brain, and to speaker preferences due to psycholinguistic processing constraints (Hawkins 2004). In fact the verb behaves like an axis around which other constituents gravitate, so this seems to obey a strategy to organize new and old information in the sentence (see for example Silva-Corvalán 1982 and Prince 1981 for a discussion on information structure). Subjects show a tendency to occur in the postverbal domain when it is not already occupied, to such a degree that the unmarked preverbal position for subjects is reversed to reach almost 70% of postverbal occurrences overall (see table 3). In the regression analysis carried out using Goldvarb, the factor or value with the highest probability of triggering a postverbal subject is in fact the absence of a co-occurring postverbal XP, with a weight of .91, followed by unaccusatives, with .80 (see table 7). But what determines if a co-occurring XP will be postverbal or preverbal? There is abundant literature suggesting that constituents tend to be organized in a way that the older elements will precede newer information (Silva-Corvalán 1982 and Prince 1981).

Relative clauses have been shown to favor postverbal subjects too (.76), becoming the fourth most important factor or value in the model, just behind presence of preverbal XP (.79), unaccusativity (.80), and absence of



postverbal XP (.91), as table 7 indicates. This observation again is supported by previous research, as mentioned in section 3.3.

Probably the most surprising result is the lack of interaction between animacy and verb type. Table 6 clearly shows that there is a statistically significant difference between the two factors ( $p < 0.05$ ) to such degree that sentences with a transitive verb, for example, are more frequently associated with inanimate subjects. How can this be explained? A close analysis of the transitive sentences included in the corpus sample suggests that there might exist a different category of subject animacy, let's call it "eventive", which is very frequent with the transitive verb *empujar* "push", but almost non-existing with the intransitives *llegar* "arrive" and *trabajar* "work". The sentences in (4) are part of the corpus sample used here and were extracted from CREA.

- (4) a. *La búsqueda de remedios más naturales empuja a los enfermos hacia distintos enfoques curativos.*  
 "The search for more natural remedies pushes the sick towards different healing approaches".
- b. *La dimisión de Guterres empuja a Portugal a unas elecciones generales anticipadas.*  
 "Guterres" resignation pushes Portugal to early elections".

The two sentences in (4) were coded as inanimate, since their subject does not correspond to a human entity, although their subject is not an inanimate object either, but rather an event or a surrounding circumstance in many cases. It is still striking that only 37.7% of all subjects with the transitive verb *push* are animate. However, when transitive verbs are omitted from the cross tabulation and only intransitives are included, then the effect of animacy becomes relevant (table 8). This statistical significance is derived from the facts observed in table 6, where unergatives are shown to have a very high percentage of animate subjects (83.8%), while unaccusatives favor inanimate subjects (63%). Thus animacy partially interacts with verb type, but only when comparing intransitives.

It is worth noting, nevertheless, that running a binomial logistic regression including only unaccusative and unergative verbs stills includes all the factor groups except for animacy which is again excluded from the best model. This may be surprising especially because at level 1 of the stepping up process all the factor groups are shown to be statistically significant when transitives are excluded.

		Subject animacy		Total	
		Animate	Inanimate		
Subject Position	Postverbal	Count	42	52	94
		Column %	24.1	47.7	33.2
	Preverbal	Count	132	57	189
		Column %	75.9	52.3	66.8
Total		Count	174	109	283
		Column %	100%	100%	100%
		Value	df	p value	
Pearson's Chi-Square		16.782	1	.0001	

**Table 8.** Subject position by Subject animacy (only unaccusatives and unergatives)

The data presented here makes obvious that animacy alone is not a good predictor of subject position. The typological research carried out by Comrie (1989), *inter alia*, presents a series of hierarchies, such as the animacy hierarchy, that interact with each other to determine how constituents will map into syntax in a given language. If Comrie's hypothesis were correct, then the sentences in (4) should have an animate subject. Multiple other hierarchies have been proposed in the literature to explain the mapping of constituents into syntax, including person, referentiality, definiteness, and thematic hierarchies. From a semantic perspective, Dowty 1991 postulated the existence of different properties related to agenthood (proto-agent properties or entailments) that determine if an argument will surface as a subject or an object. The notion of animacy is easily related to some of these properties associated by Dowty with agenthood, and therefore with subjecthood, such as volition or sentience, but there are other proto-agent entailments, such as causation, that are not. Davis and Koenig 2000 point out that, out of Dowty's proto-agent properties, causation outranks all the others when determining subjecthood. The sentences in (4) are certainly a good example of Davis and Koenig's observation. In any case, further research would be required, for example, to analyze other types of transitive verbs and the interaction between the animacy of subjects and objects in transitive constructions.

## 6. CONCLUSION

Subject position in Spanish may be either preverbal or postverbal, although there is not free variation. This paper has analyzed five variables that influence subject position in written Spanish: (i) *verb type*, (ii) *presence of a co-occurring preverbal XP*, (iii) *presence of a co-occurring postverbal XP*, (iv) *type of clause*, and (v) *subject animacy*. All of these factors were proven to have a statistically significant influence on the occurrence of postverbal subjects, except for subject animacy. Verb type and the presence of a co-occurring postverbal XP are the most significant factor groups when triggering the occurrence of a postverbal subject, followed in importance by the presence of a preverbal XP, and clause type respectively.

Regarding verb types, which could be transitive, unergative, or unaccusative, unaccusatives are the most likely trigger of subject inversion within this factor group, followed by unergatives, while transitives have an almost categorical preference for preverbal subjects.

But the value or factor that is most likely to trigger postverbal subjects is the absence of co-occurring XPs in the postverbal domain, i.e., when there is no other XP following the verb. The presence of a co-occurring preverbal XP is also a strong trigger of subject inversion. Finally, within the category of clause type (with three possible values: relative, main, and completive), relative clauses are associated with a high percentage of postverbal subjects.

It was surprising though that the regression analysis used to analyze factor weights discarded subject animacy and that the subject of transitives is more frequently inanimate. Further research is required to analyze this phenomenon, including more transitive verbs in the corpus.

The corpus used for the present research is made up of written sentences, because they are longer than oral texts, and therefore provide an opportunity for higher positional variation. Nevertheless, it would be very interesting to replicate this study using oral texts.

This research complements the existing studies that have focused on argument structure, such as those that analyze the important role that information structure plays on the position of subjects in Spanish. It also provides quantitative evidence against the assumption that the notions of subjecthood and animacy are univocally related, since the transitive verb *push* used in the present analysis has a significantly higher percentage of inanimate subjects. More research in this area would be beneficial though, including more transitive verbs into the analysis.

Finally, this paper has shown that it is possible to analyze syntactic theories, such as the Unaccusative Hypothesis, by using the variationist approach and statistical methods to further corroborate the validity of the theory.



## REFERENCES

- Ariel, Mira (1990): *Accessing noun-phrase Antecedents*. London: Routledge.
- Bel, Aurora (2003): "The syntax of subjects in the acquisition of Spanish and Catalan". *Probus: International journal of Latin and Romance linguistics* 15 (81), pp. 1-26.
- Burzio, Luigi (1981): *Intransitive verbs and italian auxiliaries*. Cambridge, MA: MIT PhD. dissertation.
- (1986): *Italian syntax*. Reidel: Dordrecht.
- Butt, John and Carmen Benjamin (2011): *A new reference grammar of modern Spanish*. 5<sup>th</sup> edition. London and New York: Routledge.
- Chafe, Wallace (1976): "Givenness, contrastiveness, definiteness, subjects, topics, and point of view", in Charles Li (ed.): *Subject and topic*. New York: Academic Press. pp. 25-55.
- (1987): "Cognitive constraints on information flow", in Russell S. Tomlin (ed.): *Coherence and grounding in discourse. Typological studies in language* 11. Amsterdam: John Benjamins. pp. 21-52.
- (1994): *Discourse, consciousness, and time*. Chicago/London: University of Chicago Press.
- Comrie, Bernard (1989): *Language universals and linguistic typology*. Chicago: The University of Chicago Press.
- Davis, Anthony R. and Jean-Pierre Koenig (2000): "Linking as constraints on word classes in a hierarchical lexicon". *Language* 76, pp. 56-91.  
<http://dx.doi.org/10.1353/lan.2000.0068>  
<http://dx.doi.org/10.2307/417393>
- Díaz-Campos, Manuel (2001): *La posición de la frase nominal sujeto respecto al verbo: un estudio del habla infantil caraqueña*. Caracas: Fondo Editorial de la Facultad de Humanidades y Educación, Universidad Central de Venezuela.

- Dowty, David (1991): "Thematic proto-roles and argument selection", *Language* 67, pp.547-619.  
<http://dx.doi.org/10.2307/415037>  
<http://dx.doi.org/10.1353/lan.1991.0021>
- Firbas, Jan (1964): "On defining the theme in functional sentence analysis", *Travaux Linguistiques de Prague* 1, pp. 267-280.
- Gundel, Jeanette, Nancy Hedberg and Ron Zacharski (1993): "Cognitive status and the form of referring expressions in discourse", *Language* 69, pp. 274-307.  
<http://dx.doi.org/10.2307/416535>
- Gutiérrez-Bravo, Rodrigo (2003): "Subject inversion in Spanish relative clauses: a case of prosody-induced word order variation without narrow focus", in Twan Geerts, Ivo van Ginneken and Haike Jacobs (eds.): *Romance languages and linguistic theory 2003*. Amsterdam: John Benjamins, pp. 115-138.
- Hawkins, John A. (2004): *Efficiency and complexity in grammars*. Oxford: Oxford University Press.  
<http://dx.doi.org/10.1093/acprof:oso/9780199252695.001.0001>
- Kahn, Jason M. and Jennifer E. Arnold (2012): "A Processing-Centered Look at the Contribution of Givenness to Durational Reduction". *Journal of Memory and Language* 67, pp. 311-325.  
<http://dx.doi.org/10.1016/j.jml.2012.07.002>
- Lambrecht, Knud (1994): *Information Structure and Sentence Form. Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.  
<http://dx.doi.org/10.1017/CBO9780511620607>
- Lipski, John M. (1977): "Preposed Subjects in Questions: Some Considerations", *Hispania* 60 (1), pp. 61-67.  
<http://dx.doi.org/10.2307/340393>
- Mayoral Hernández, Roberto (2007): "A variation study of verb types and subject position: Verbs of light and sound emission", in José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Viviane Déprez and María José Cabrera (eds.): *Romance Linguistics 2006, Selected papers from the 36th Linguistic Symposium on Romance Languages*. Amsterdam / Philadelphia: John Benjamins, pp. 213-226.

- (2008): *The locative alternation: Unaccusative constructions and subject position*. Los Angeles, CA: University of Southern California dissertation.
  - (2010): “The locative alternation: On the symmetry between verbal and prepositional locative paradigms”. *Probus: International Journal of Latin and Romance Linguistics*, 22 (2), pp. 211-239
- Ocampo, Francisco (2003): “On the notion of focus in spoken Spanish: An empirical approach”, in Paula Kempchinsky and Carlos-Eduardo Piñeros (eds.): *Theory, Practice, and Acquisition*. Sommerville: Cascadilla Press, pp. 207-226.
- Orozco, Rafael and Gregory Guy (2008): “El uso variable de los pronombres sujetos: ¿qué pasa en la costa Caribe colombiana?”, in Maurice Westmoreland and Juan Antonio Thomas (eds.): *Selected proceedings of the 4<sup>th</sup> workshop on Spanish sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project, pp. 70-80.
- Ortiz López, Luis A. (2010): “El español del Caribe: orden de palabras a la luz de la interfaz léxico-sintáctica y sintáctico-pragmática”, *Revista Internacional de Lingüística Iberoamericana*, 7.2 (14), pp. 75-93.
- Poplack, Shana. (1980): “Deletion and disambiguation in Puerto Rican Spanish”, *Language* 56, pp. 371-385.  
<http://dx.doi.org/10.2307/413761>  
<http://dx.doi.org/10.1353/lan.1980.0033>
- Perlmutter, David (1978): “Impersonal passives and the unaccusative hypothesis”, in Jeri J. Jaeger et al, (eds.): *Proceedings of the 4th annual meeting of the Berkeley Linguistics Society*, BLS: Berkeley, CA, pp. 157-189.
- Prince, Ellen F. (1981): “Toward a taxonomy of given-new information”, in P. Cole (ed.): *Radical pragmatics*. New York: Academic Press, pp. 223-256.
- Real Academia Española, Banco de datos (CREA) online, *Corpus de Referencia del Español Actual*.  
<http://www.rae.es>.
- Silva-Corvalán, Carmen (1982.): “Subject expression and placement in Mexican American Spanish”, in Jon mastae and Lucía Elías-Olivares (eds.): *Spanish in the United States: Sociolinguistic aspects*. Cambridge/New York: Cambridge University Press, pp. 93-120.



- (2003): “Otra mirada a la expresión del sujeto como variable sintáctica”, in Francisco Moreno Fernández et al. (eds.): *Lengua, variación y contexto: Estudios dedicados a Humberto López Morales*. Madrid: Arco/Libros, pp. 849-60.
- Schütze, Carson T. (1996): *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.
- Torrego, Esther (1989): “Unergative unaccusative alternations”, *MIT Working Papers in Linguistics* 10, pp. 253-272.
- Zubizarreta, María Luisa (1998): *Prosody, Focus, and Word Order*. Linguistic Inquiry Monograph 33. Cambridge, MA: MIT Press.
- and Eunjeong Oh (2007): *On the syntactic composition of manner and motion, linguistic inquiry monograph*. Cambridge, MA: MIT Press.

## Uso del sujeto y del objeto directo en el habla de once ciudades hispanohablantes

Mercedes Sedano y Paola Bentivoglio  
*Universidad Central de Venezuela*

**Abstract:** We present an exploratory quantitative analysis of the form and position of the main constituents —subject (S) and direct object (OD)— in a corpus of two speakers, a man and a woman, middle aged and with higher education, belonging to five cities from Spain and six from the New World, respectively. We analyzed 2,200 words of twenty two speakers recorded for the “Proyecto del Estudio Sociolingüístico de España y de América (PRESEEA)”. Our goal was to ascertain whether basic word order varies substantially from one place to another. The results, though limited because of the small number of speakers and correspondent tokens, show that: i) S’s tend to be preverbal and OD’s postverbal; ii) frequency of null S’s is higher; iii) S-clauses tend to be postverbal; iv) OD’s, independently of their form, are almost categorically postponed; v) the pronominal subject most frequently used is *yo*; and vi) there are no dialect or gender-related differences. Our results confirm those of previous extended research on word order both in Peninsular and American dialects.

**Keywords:** Spoken Spanish, word order, subject, direct object.



## 1. INTRODUCCIÓN\*

El objetivo del presente trabajo es determinar cuantitativamente el uso en español del sujeto (S) y del objeto directo (OD) y, en particular, su posición con respecto al verbo, en un corpus de habla proveniente de once ciudades hispanohablantes.

Desde el punto de vista tipológico, el español se suele describir como una lengua *pro-drop*, es decir, con un sujeto que puede ser implícito o explícito. Otra característica del español es su orden (S)VO flexible; esto significa que, aunque el sujeto suele aparecer antes del verbo, otros órdenes son también posibles aunque menos frecuentes (Greenberg 1963, Bentivoglio y D'Introno 1989).

En las cláusulas informativas normales, no marcadas<sup>1</sup>, el primer constituyente suele ser el sujeto, que normalmente contiene información conocida y corresponde a un referente con las propiedades de ser determinado, agentivo y humano o, al menos, animado (Comrie 1981, Tomlin 1985). En circunstancias habituales, del sujeto clausular situado en posición preverbal se avanza linealmente, a través del verbo, hasta la información que se supone menos conocida por el interlocutor (Silva-Corvalán 2001).

Aunque actualmente existen muchos trabajos que buscan esclarecer los distintos factores que condicionan la posición del S y del OD, nosotras queremos simplemente presentar un análisis exploratorio del uso de las oraciones del español informal hablado respecto de la forma y posición del sujeto, implícito o explícito, independientemente del verbo y de los complementos con los que estén construidas, así como del objeto directo, cuando este está presente ¿Cómo se organizan esas oraciones? ¿Cuál es la proporción de sujetos implícitos con respecto a sujetos explícitos? ¿Qué forma adoptan el S y el OD: pronombre, frase nominal, cláusula subordinada? ¿Hay alguna relación entre la forma del S y del OD y su aparición preverbal o postverbal? ¿Hay diferencias que podrían considerarse dialectales? Aunque esta última hipótesis parece haber sido desechada, cabe comprobar si en el corpus de las ciudades donde se elide o suaviza la /-s/ al final de palabra hay un incremento de sujetos expresos para evitar la ambigüedad de algunas formas verbales (*¿sabe o sabe(s)? ¿baila o baila(s)?*, (Cameron 1992).

---

\* Queremos agradecer a los árbitros anónimos sus observaciones y sugerencias, que nos han ayudado a mejorar el presente texto. Huelga añadir que los errores restantes son de nuestra exclusiva responsabilidad.

<sup>1</sup> El orden marcado o no marcado de una cláusula en español se correlaciona también con la tipología verbal, según lo afirman varios lingüistas (ver Bentivoglio 1989a), lo que se comprueba, por ejemplo, en las cláusulas con verbos existenciales, en las que el orden no marcado es VS y no SV.

## 2. ANTECEDENTES DE ESTUDIO

En las gramáticas tradicionales se consideraba, como hace notar la *Nueva gramática de la lengua española* (NGRALE 2009 §40.1s), que el orden de palabras en una oración era un recurso retórico para lograr determinados matices expresivos. Dentro de esa tónica, se diferenciaba el “orden lineal” (sujeto-verbo-complemento directo-complemento indirecto-circunstanciales) del “orden envolvente”, que suponía una alteración del orden lineal (Gili Gaya 1961 §70, Real Academia Española 1973 §3.7.3).

En la segunda mitad del siglo pasado se comenzó a dar una gran importancia al orden básico de palabras porque se observó que ese orden se correlacionaba con muchos otros aspectos de una lengua, por ejemplo, si esta tenía pre- o posposiciones o si el antecedente de una cláusula relativa precedía a la cláusula o aparecía después (Greenberg 1963; Croft 1990). Para estudiar el orden básico de las palabras o, más precisamente, el orden de los constituyentes en una oración<sup>2</sup>, se consideró fundamental establecer la posición del sujeto y/o del objeto directo con respecto al verbo (V)<sup>3</sup>. Entre los trabajos pioneros sobre orden de constituyentes en nuestra lengua, están los de Bolinger (1954-55), Hatcher (1956), y Contreras (1976). El artículo de Bolinger sigue siendo fundamental aún hoy, a casi sesenta años de su publicación en el *Boletín de Filología* de la Universidad de Chile dedicado a Rodolfo Oroz. La investigación de Hatcher se basa fundamentalmente en el español escrito, mientras que Contreras, siguiendo el modelo generativista de la época, usa la introspección para tratar de formalizar reglas específicas.

Los primeros estudios basados en corpus de habla sobre el orden de los constituyentes oracionales del español hablado son los de Silva-Corvalán (1977, 1982, 1983), Bentivoglio y D’Introno (1978, 1989), Morales (1982, 1986), Bentivoglio y Weber (1986), y Bentivoglio (1988, 1989a), todos basados en muestras de español americano<sup>4</sup>. Por su parte,

<sup>2</sup> De ahora en adelante, y siguiendo la apropiada propuesta de diversos autores (Silva-Corvalán 2001, Ocampo, en prensa), nos referiremos al orden de los constituyentes.

<sup>3</sup> Dryer (1992) establece tipologías de lenguas basándose exclusivamente en la posición del OD con respecto al verbo.

<sup>4</sup> Las investigaciones de Bolinger (1954-55), Hatcher (1956), Contreras (1976), así como las de Silva-Corvalán (1977 y 1982), Morales (1982), Bentivoglio y D’Introno (1978, 1989), Bentivoglio y Weber (1986), y Bentivoglio (1988 y 1989b), están reseñadas más detalladamente en Bentivoglio (1989a).

Delbecque (1979, 1987 y 1991) aborda el mismo tema en un corpus de español peninsular escrito.

Hoy en día hay numerosas investigaciones sobre el tema, muchas de las cuales se han centrado en la posición preverbal o postverbal del sujeto, pero desde distintas perspectivas. Desde el punto de vista del *tipo de oración*, unos investigadores han estudiado la posición del sujeto en las oraciones transitivas, otros, en las intransitivas (Alfonso Vega y Melis 2010) o en las copulativas (Bentivoglio 1989b, Ocampo 2005), y también en las oraciones con dos o tres constituyentes (Ocampo, en prensa). Desde el punto de vista del *tipo de sujeto*, en unos trabajos se estudia cualquier tipo de sujetos, mientras que en otros se analizan los sujetos pronominales en general (Fernández Soriano 1999, Posio 2012, Serrano 2013), o bien un determinado pronombre personal, que suele ser el correspondiente a la 1ª persona *yo* (Cantero 1976, Bentivoglio 1987, Blanco Canales 1999, Samper Padilla, Hernández Cabrera y Troya Déniz 2006, Travis 2007, y Pérez Brabandere 2010), pero también *yo* y *tú* (Davidson 1996), *usted/ustedes* (Serrano 2013), los de 3ª persona (Dumont 2006, Quijada 2006, Balasch 2008, Martínez Lara 2012), e incluso los de 1ª, 2ª y 3ª persona (Barrenechea y Alonso 1977, Silva-Corvalán 1977, Enríquez 1984 y 1986, Morales 1986 y 1987, Miró Vera y de Pineda 1990, Cameron 1995, Lu 1997, Hurtado 2001, Cameron y Flores Ferrán 2004). Muchas de las investigaciones mencionadas se han llevado a cabo siguiendo los principios de la sociolingüística variacionista.

Los resultados de las investigaciones realizadas hasta ahora en español indican, por un lado, que hay restricciones para el uso del sujeto implícito (Silva-Corvalán, en prensa) y, por el otro, que, en los casos de sujeto explícito, su anteposición al verbo es mucho más frecuente que su posposición; entre los factores que influyen en la posición del sujeto se han señalado el tipo de información (conocida o nueva) que transmite la entidad sujeto (Bentivoglio 1993, Silva-Corvalán 2001, Ocampo 2009), su topicalidad o tematicidad (Morales 1983; Ocampo 2009 y en prensa), o bien su carácter [ $\pm$ determinado], [ $\pm$ agentivo], [ $\pm$ humano] o [ $\pm$ pronominal] (Morales 1983; Delbecque 2005). También el tipo de verbo: cognición, existencia, etc. (Bentivoglio 1987, 1989a y 2003, Silva-Corvalán 2001, Posio 2012), su carácter [ $\pm$ transitivo]<sup>5</sup> (Delbecque 2005, Serrano 2013) o incluso el número de constituyentes de la oración (Alfonso Vega y Melis 2010; Ocampo en prensa).

<sup>5</sup> Vázquez Rozas (2004) hace notar que los verbos transitivos más frecuentes son los de baja transitividad.

Hasta donde llegan nuestros conocimientos, hay menos estudios sobre el uso del OD en español que sobre el uso del sujeto. Sin embargo, todos los trabajos existentes coinciden en que, en circunstancias normales, y con excepción de los clíticos acusativos preverbales, los ODs están pospuestos al verbo, sobre todo cuando corresponden a una frase nominal (FN). Su aparición preverbal es excepcional (Morales 1983, Bentivoglio y D’Introno 1989, Bogard 2010), y se asocia frecuentemente a construcciones del tipo *El libro lo compré ayer*, denominadas de *dislocación/tematización a la izquierda* (Hidalgo Downing 2003, Martín Butragueño 2010, Sedano 2012), en las cuales el OD preverbal (*El libro*) suele concordar con un pronombre correferencial situado en el interior de la cláusula correspondiente (*lo*). Silva-Corvalán (1984) matiza las distintas posibilidades de aparición del OD preverbal según la combinación de los rasgos [ $\pm$ nuevo] y [ $\pm$ contrastivo] y muestra cuatro tipos de funciones que puede realizar el OD antepuesto: i) enlace textual, ii) foco contrastivo, iii) contrario a lo esperado, y iv) complemento focal.

Los resultados de las numerosas investigaciones realizadas sobre el orden de los constituyentes en español indican que los factores condicionantes de ese orden pueden ser cognitivo-sintácticos, semánticos y pragmático-discursivos; entre estos últimos puede mencionarse el tipo de discurso en el que se enmarca el texto analizado (Vázquez Rozas y García Miguel 2006). No parece que los factores extralingüísticos de la edad, el sexo o el nivel socioeconómico de los hablantes influyan en la posición de los constituyentes (Bentivoglio 1987: 65; Silva-Corvalán 2001: 155), aun cuando en algunas investigaciones se anotan tendencias diferentes, sobre todo debidas al sexo (ver Serrano 2013).

### 3. METODOLOGÍA

Hemos analizado todos los casos de sujeto, no importa si la oración es transitiva o intransitiva, y los casos de OD (excepto los indicados en los puntos iv y viii en el párrafo siguiente) en un corpus de transcripciones de habla del “Proyecto para el Estudio Sociolingüístico del Español de España y de América” (PRESEEA; Moreno Fernández 1997, 2005). Las ciudades cuyos datos hemos estudiado son, por orden alfabético, las siguientes: Alcalá de Henares, Buenos Aires, Caracas, Granada, La Habana, Lima, Las Palmas de Gran Canaria, Madrid, Ciudad de México, Santiago de Chile, y Santiago de

Compostela<sup>6</sup>. Hay por tanto cinco ciudades españolas y seis latinoamericanas. En cada ciudad se analizan 2200 palabras<sup>7</sup> del habla de un hombre y otras tantas de una mujer con edades comprendidas entre 35 y 54 años y con 15 años de escolarización o más<sup>8</sup>.

Hemos seleccionado todas las oraciones declarativas, afirmativas y negativas, que tienen un sujeto, explícito o implícito, no importa si están construidas con un objeto directo o no. Señalamos a continuación algunas decisiones que hemos tomado en casos que podrían ser poco claros en cuanto a su inclusión o exclusión del análisis:

- i) Considerar que solo hay un sujeto en cada perífrasis verbal (*podía hacerlo, estaba cantando, iba a terminar de escribir, ha de hacerlo*; NGRALE 2009 § 28.1. y ss.).
- ii) Analizar el pronombre *se* como sujeto cuando aparece en una cláusula impersonal intransitiva (*Se sube por esta calle*)<sup>9</sup>, pero no hacerlo cuando *se* aparece en construcciones como *se dividió la casa*, debido al hecho de que no queda claro si *se* representa a un verdadero sujeto impersonal (*se* = ‘alguien’ *dividió la casa*) o bien *se* trata de una construcción pasiva refleja (‘la casa fue dividida’)<sup>10</sup>.
- iii) Tomar en cuenta solo las cláusulas interrogativas, exclamativas y exhortativas<sup>11</sup> que funcionan en bloque como OD (*decidiría qué*

<sup>6</sup> Deseamos expresar nuestro agradecimiento a los coordinadores del proyecto PRESEEA en las ciudades que hemos seleccionado para nuestro estudio por su gentil envío de las transliteraciones de habla correspondientes a esas ciudades.

<sup>7</sup> Para seleccionar esas 2200 palabras fue necesario eliminar antes de las transcripciones correspondiente tanto las palabras del entrevistador o entrevistadores como las indicaciones metalingüísticas: silencio, risas, ruidos, ininteligible, etc.

<sup>8</sup> En el apéndice se muestran los códigos que identifican las ciudades, así como los que identifican individualmente a cada hablante; estos últimos códigos fueron ofrecidos gentilmente por los coordinadores de PRESEEA en cada una de las ciudades estudiadas. Para no alargar innecesariamente la presentación de los datos, en cada cuadro y en cada ejemplo aparecerá simplemente el nombre de la ciudad y el sexo del hablante: H si es hombre y M si es mujer. Si los lectores quieren más información, podrán consultar el apéndice. No se hace referencia ni a la edad de los hablantes ni a su nivel socioeducativo, porque, como ya lo hemos mencionado, todos ellos tienen entre 35 y 54 años, y han recibido 15 años de escolarización o más.

<sup>9</sup> En NGRALE (2009 §41.10g) se considera que el clítico *se*, en las oraciones claramente impersonales con *se* (*se* = ‘alguien’ *salió muy temprano*), es “la marca o el exponente de un sujeto tácito”.

<sup>10</sup> Sobre este tema, puede consultarse NGRALE (2009 §41.10g., 41.1a. y 41.12.).

<sup>11</sup> Sobre la diferencia de orden entre los constituyentes de las oraciones declarativas y las no declarativas puede consultarse, por ejemplo, Aijón Oliva y Serrano (2010) y Serrano (2013), donde se hace notar que en las interrogativas aumentan los casos de posposición.



- era lo que iba a estudiar**); en estos últimos casos no se analiza el funcionamiento interno de la cláusula subordinada.
- iv) Descartar los casos de OD implícito, ya que no siempre resulta fácil distinguir un verdadero OD implícito (*Yo quiero ir pero él no quiere*) del uso absoluto de un verbo transitivo, como en **leía mucho** (NGRALE §34.4b y ss.)<sup>12</sup>.
  - v) En las oraciones de dislocación a la izquierda (El libro sí **lo había leído**) y de dislocación a la derecha (Sí **lo había leído**, el libro), y debido al carácter extraclausal que se suele atribuir al constituyente dislocado (*el libro*)<sup>13</sup>, no analizar ese constituyente sino el que, siendo correferencial con él, cumple una verdadera función sintáctica en el interior de la cláusula correspondiente (**lo**).
  - vi) Excluir los marcadores del discurso: *digamos, imagínate, o sea, a ver, pero vamos, es decir*, etc.
  - vii) No analizar las oraciones hendidas (*Fue él quien saludó*), pseudohendidas inversas (*Él fue quien saludó*) y pseudohendidas (*Quien saludó fue él*), porque su función focalizadora hace que ofrezcan características particulares en cuanto al orden de los constituyentes (NGRALE 2009 §40.1a-e).
  - viii) No considerar el argumento asociado a los verbos *haber* y *hacer* en su uso impersonal por tratarse de verbos con características particulares.

#### 4. RESULTADOS DEL ANÁLISIS

La presente sección se divide en cuatro apartados: en el primero, se analizan los resultados correspondientes al sujeto; en el segundo, se hace otro tanto con el OD; en el tercero, se ofrecen dos cuadros con los resultados generales sobre los usos del sujeto y del OD; y, finalmente, en el cuarto apartado se analizan algunos resultados extralingüísticos.

<sup>12</sup> Para ilustrar la subjetividad que puede rodear la selección de un OD implícito, vale la pena señalar, por ejemplo, que mientras Morales (1983: 483) sólo encuentra 2 casos en una muestra de 1636 oraciones (0,12%), Posio (2012) registra un 21% de las 1660 cláusulas estudiadas.

<sup>13</sup> Véanse, por ejemplo, Lambrecht (2001), Hidalgo Downing (2003), Martín Butragueño (2010), y Sedano (2012).

Cada apartado empieza con la presentación de los ejemplos que ilustran los casos analizados. Los parámetros en los que estructuralmente es posible que el S o el OD aparezca antepuesto o pospuesto al verbo se esclarecen con dos pares de ejemplos –uno relativo a la anteposición y otro a la posposición– que llevan el mismo número, pero el segundo, el de la posposición, tiene un apóstrofo. Los casos en que estructuralmente solo es posible una de las dos posiciones se ilustran con un solo par de ejemplos. Al final de cada ejemplo se ofrece entre paréntesis la identificación de la ciudad y del hablante.

Para facilitar la lectura de los ejemplos, tanto el S como el OD aparecen en cursiva, excepto cuando el S es implícito y está representado por el signo Ø. Los verbos de las cláusulas donde están el S y el OD se resaltan en negritas.

#### 4.1. El sujeto

##### 4.1.1. El sujeto implícito

Los sujetos representados por el símbolo Ø son los que se expresan exclusivamente por medio de los morfemas de persona y número en la forma verbal conjugada, o bien están del todo ausentes en las formas no conjugadas del verbo (infinitivos y gerundios). Ambos tipos se ilustran en (1a) y (1b), respectivamente:

1. /sujeto implícito Ø/
  - a. Ø **Haces** el arroz, Ø lo **pones** a Ø **refrescar** (La Habana H)
  - b. la sensación de Ø **ir** con ropa de invierno Ø **pasando** calor (Madrid-H)

Estudiamos estos sujetos para comprobar si, efectivamente, son mucho más numerosos que los sujetos explícitos, como se señala en todos los estudios reseñados sobre el tema. En el cuadro 1 se ofrecen los resultados de la distribución de los sujetos implícitos en los dos hablantes de cada ciudad; para cada uno de ellos se indican las frecuencias absolutas y relativas de esos sujetos con verbo conjugado (2ª y 3ª columnas) y verbo no conjugado (4ª y 5ª); en la 6ª y 7ª columnas aparecen el total y el porcentaje de todos los sujetos implícitos encontrados en el corpus.

	Con verbo conjugado	%	Con verbo no conjugado	%	Total	%
Alcalá H	188	95	10	5	198	5,89%
Alcalá M	126	94	6	6	134	3,98%
Buenos Aires H	142	96	6	4	148	4,40%
Buenos Aires M	151	99	1	1	152	4,52%
Caracas H	160	94	9	6	169	5,02%
Caracas M	137	95	7	5	144	4,28%
Granada H	150	95	7	5	157	4,67%
Granada M	155	95	8	5	163	4,85%
La Habana H	142	93	10	7	152	4,52%
La Habana M	159	95	8	5	167	4,96%
Las Palmas H	123	92	10	8	133	3,95%
Las Palmas M	158	97	5	3	163	4,85%
Lima H	143	92	13	8	156	4,64%
Lima M	157	96	6	4	163	4,85%
Madrid H	136	88	19	12	155	4,61%
Madrid M	152	94	9	6	161	4,79%
México H	134	96	6	4	140	4,16%
México M	102	78	28	22	130	3,86%
Santiago, Chile H	129	92	11	8	140	4,16%
Santiago; Chile M	176	96	8	4	184	5,47%
Santiago de Com. H	111	94	7	6	118	3,51%
Santiago de Com. M	135	99	2	1	137	4,07%
Total	3225	96	139	4	3364	100,00%

Cuadro 1. Distribución de los sujetos implícitos (Ø) con verbo conjugado y no conjugado

El cuadro 1 muestra que la cantidad de sujetos implícitos con verbos conjugados excede en mucho a la de sujetos implícitos con verbos no conjugados: 96% vs. 4%; el promedio de los sujetos implícitos con verbo no conjugado se mantiene muy similar en todos los hablantes (entre 3% y 8%), con un mínimo de 1% en las dos mujeres de Buenos Aires y de Santiago de Compostela, respectivamente, y un máximo de 22% en la mujer de Ciudad de México. La distribución total de los sujetos implícitos es bastante parecida en los hablantes, hombres y mujeres, de las once ciudades: los porcentajes varían entre un mínimo de 3,51% en el hombre de Santiago de Compostela y un máximo de 5,89% en el hombre de Alcalá de Henares.

#### 4.1.2. Los pronombres

En esta sección presentamos las distintas formas del sujeto antepuesto y pospuesto al verbo, menos los casos ilustrados en el ejemplo (2), con un sujeto *se* que en el corpus siempre está antepuesto:

2. /sujeto *se* en oración impersonal intransitiva/
  - a. luego pues *se jugaba* con barro (Madrid H)
  - b. *se podía salir* sin problema (Granada H)

También aparecen siempre antepuestos al verbo los sujetos que son pronombres relativos y que se ejemplifican más adelante en (8a-b).

Los S cuya forma es un pronombre personal se ilustran en los ejemplos (3-5'), empezando con los personales, que están separados en tres grupos: *yo, tú/vos*<sup>14</sup>, y otros pronombres personales tónicos (*él, usted, nosotros*, etc.); los pronombres personales van seguidos por los demostrativos en (6-6'), los cuantitativos (*algo, mucho, todo, nada*, etc.)<sup>15</sup> en (7-7'), y los relativos en (8):

3. /sujeto antepuesto: pronombre personal *yo*/
  - a. *yo me me iba* a a/ pasear todo el santo día/ (México H)
  - b. *yo me molestaba* (Caracas M)
- 3.' /sujeto pospuesto: pronombre personal *yo*/
  - a. la parte alta de Peñalolén ahí *vivo yo* (Santiago, Chile H)
  - b. me *había ido yo* a Chapultepec/ (México H)

<sup>14</sup> No hemos incluido en este parámetro el pronombre *usted/es* porque, si bien pragmáticamente corresponde a la 2ª persona, formalmente se relaciona con la 3ª, no con la 2ª. Cabe señalar además que apenas hay casos de *usted/es* en el corpus analizado.

<sup>15</sup> Se incluyen en este grupo los pronombres que la NGRALE (2009 §19.7) denomina *cuantitativos*, entre los cuales en esa gramática están también los indefinidos.

4. /sujeto antepuesto: pronombre personal *tú/vos*/
  - a. *tú* **eras** más joven y tal (Alcalá H)
  - b. escanean las fotos que *vos* **vas a ver** en la revista (Buenos Aires M)
- 4'. /sujeto pospuesto: pronombre personal *tú/vos*/
  - a. pero que te lo **tienes que organizar** *tú* (Santiago de Compostela H)
  - b. y **decías** *tú*: “¿qué es una porra?” (Santiago de Compostela M)
5. /sujeto antepuesto: otros pronombres personales tónicos/
  - a. *él* se me **acercaba** (Caracas M)
  - b. *ella* **quería** estudiar Bellas Artes (Las Palmas M)
- 5'. /sujeto pospuesto: otros pronombres personales tónicos/
  - a. ella decía “No, no; la linda **es ella**” (Las Palmas M)
  - b. era muy cerquita de de la colonia donde **vivíamos** *nosotros* (México H)
6. /sujeto antepuesto: pronombre demostrativo/
  - a. *eso* me **duele** (La Habana M)
  - b. no, mira, *aquellos* **son** mejor (Caracas M)
- 6'. /sujeto pospuesto: pronombre demostrativo/
  - a. me **gusta** *esa* (La Habana H)
  - b. ya yo tenía treinta años cuando **pasaba** *eso* (Buenos Aires H)
7. /sujeto antepuesto: pronombre cuantitativo/
  - a. *uno* en esta vida **hace** lo que puede (Caracas M)
  - b. *nadie* me **dijo** na(da) (Granada M)
- 7'. /sujeto pospuesto: pronombre cuantitativo/
  - a. **cabíamos** *todos* (Madrid H)
  - b. y nada, no **pasó** *nada* (Caracas H)
8. /sujeto antepuesto: pronombre relativo/
  - a. vivo en Peñalolén en una zona *que* se **llama** Casagrande (Santiago, Chile H)
  - b. mi madre ... había tenido una prima *que* **cantaba** (Las Palmas H)

Las frecuencias absolutas de los varios tipos de sujetos pronominales se ofrecen en el cuadro 2:

	Pronombre personal tónico							Pron. demostrat.		Pron. cuantitat.		Pron. relat.	Total
	<i>se</i>	<i>yo</i>		<i>tú/vos</i>		Otros							
	Ant.	Ant.	Pos.	Ant.	Pos.	Ant.	Pos.	Ant.	Pos.	Ant.	Pos.	Ant.	
Alcalá H	0	11	0	1	0	3	1	1	0	1	1	10	29
Alcalá M	0	16	1	0	0	1	1	2	0	3	0	10	34
Buenos Aires H	0	25	1	0	0	6	0	6	1	4	1	10	54
Buenos Aires M	0	10	0	3	0	4	0	10	0	15	0	9	51
Caracas H	0	3	0	1	0	6	0	5	0	3	3	17	38
Caracas M	0	34	0	7	0	15	0	7	0	7	0	10	80
Granada H	3	33	0	5	0	8	0	8	0	3	2	10	72
Granada M	0	16	1	1	0	5	1	7	0	2	0	8	41
La Habana H	0	11	0	4	0	7	0	6	1	2	2	17	50
La Habana M	0	4	0	2	0	2	0	6	0	9	1	18	42
Las Palmas H	0	30	2	0	0	2	0	2	0	3	0	21	60
Las Palmas M	0	17	0	4	0	21	2	2	0	4	0	5	55
Lima H	1	22	0	1	0	2	0	6	1	5	0	12	50
Lima M	1	23	0	3	0	1	0	6	0	4	0	15	53
Madrid H	2	27	1	1	0	4	0	0	0	0	2	9	46
Madrid M	0	11	0	0	0	3	1	3	0	2	0	9	29
México H	0	34	2	0	0	9	0	4	0	4	4	6	63
México M	0	11	2	0	0	2	0	0	0	1	0	12	28
Santiago Chi. H	1	20	1	0	0	7	0	6	1	4	1	14	55
Santiago Chi. M	1	10	0	0	0	2	0	2	0	0	2	6	23
Santiago Co. H	0	2	0	3	1	5	0	4	0	2	0	25	42
Santiago Co. M	0	14	0	3	1	19	0	2	0	4	0	11	54
Total	9	384	11	39	2	134	6	95	4	82	19	264	1049

**Cuadro 2.** Uso de los distintos sujetos pronominales

En el cuadro 2 se observa que los usos de sujeto pronominal más frecuentes corresponden a la 1ª persona *yo*, que generalmente antecede al verbo (384 casos de anteposición = 97% vs. 11 de posposición = 3%). El elevado uso de este pronombre podría atribuirse al egocentrismo del lenguaje, sobre todo el que se da en la conversación<sup>16</sup>. Hay pocos casos de sujetos de 2ª persona (39 antepuestos = 95% y 2 pospuestos = 5%), y estos casos, por lo general, no se refieren al interlocutor sino que se interpretan en sentido genérico ('cualquier persona'). Esto es esperable ya que los entrevistadores dejan que el hablante se exprese libremente y lo interrumpen lo menos posible.

Los restantes pronombres personales son bastante numerosos: los 134 casos de anteposición (96%) superan ampliamente a los 6 de posposición (4%). Son, sin embargo, menos frecuentes que los relativos (264 casos, todos antepuestos). Dentro de los pronombres demostrativos, hay 95 antepuestos (96%) y 4 pospuestos (4%) y, dentro de los cuantitativos, 82 antepuestos (81%) y 19 pospuestos (19%).

Llama la atención que los porcentajes de posposición de los pronombres sujeto no superen el 6%, salvo cuando se trata de *yo* y de los cuantitativos, donde la posposición alcanza el 11% y el 19%, respectivamente. Es posible que, en el caso de *yo*, esto se deba a razones pragmáticas –por ejemplo, *creo/digo/pienso yo*–, mientras que, en el caso de los indeterminados, podría ser justamente su carácter indeterminado (véanse los ejemplos 7 y 7') lo que contribuye a que no se perciban como sujetos prototípicos (determinados y humanos o, al menos, animados) y que, en consecuencia, se asimilen en posición al OD.

#### 4.1.3. Las frases nominales léxicas y las cláusulas

En este apartado incluimos dos tipos de sujetos: el primero comprende las frases nominales léxicas y, el segundo, los sujetos clausulares. Las FNs léxicas pueden ser determinadas (ejs. 9 y 9') o indeterminadas (ejs. 10 y 10')<sup>17</sup>; tanto unas como otras suelen estar modificadas, entre otros constituyentes, por adjetivos, frases preposicionales, o por una cláusula

<sup>16</sup> Morales (1983: 480) también hace notar el carácter egocéntrico del discurso al estudiar en su corpus el empleo de *yo*. Igual sucede con Vázquez Rozas y García Miguel (2006), quienes observan que, aunque el sujeto *yo* solo es superado en frecuencia por la 3ª persona del sing., *yo* se emplea porcentualmente mucho más que cualquier otro sujeto cuando se trata de verbos de proceso mental (*creer, saber ...*).

<sup>17</sup> Para analizar una FN como determinada o como indeterminada, hemos seguido los criterios establecidos por Sedano (2011: 454).

relativa. Dentro de los sujetos clausulares están las cláusulas sustantivas y las frases infinitivas (11 y 11'); también se incluyen en este grupo las cláusulas relativas sin antecedente (ver ej. 12) que funcionan como sujeto (NGRALE 2009 §33.10h)<sup>18</sup>:

9. /sujeto antepuesto: FN determinada/
  - a. *el nivel de desenfreno* **es** bastante alto (Santiago, Chile H)
  - b. *su padre* **era** médico (Santiago de Compostela M)
- 9'. /sujeto pospuesto: FN determinada/
  - a. o sea **se quedaba** *la casa* sin piscina (Granada M)
  - b. a mí me **gusta** más *el centro* (Alcalá M)
10. /sujeto antepuesto: FN indeterminada/
  - a. *un paisaje para mí lindo* **es** un paisaje así invernal, con frío (Lima M)
  - b. con las ventanas abiertas *algo de fresco* te **tiene que entrar** (Alcalá H)
- 10'. /sujeto pospuesto: FN indeterminada/
  - a. en el edificio **queda** *un abasto* abajo (Caracas H)
  - b. Ahí me **empezaron a llamar** *distintos fotógrafos* (Buenos Aires H)
11. /sujeto antepuesto: cláusula sustantiva/
  - a. verlos comer me **encanta** (La Habana H)
  - b. *ir en contra* **era** más bien lo raro (Lima M)
- 11'. /sujeto pospuesto: cláusula sustantiva/
  - a. y también me **da** cosa que esté ahí la criatura (Granada M)
  - b. me **gustaba** mucho *salir* sin el permiso de mi mamá (México H)
12. /sujeto antepuesto: cláusula relativa/
  - a. ahí *lo que yo gano de esas cosas* no **es** demasiado (Buenos Aires H)
  - b. son tiempos que incluso *quien te está hablando* no los **ha vivido** (Madrid H)

<sup>18</sup> Consideramos que una cláusula relativa sin antecedente léxico puede funcionar toda ella como S o como OD; por el contrario, cuando la relativa tiene antecedente léxico, es el pronombre relativo que representa a ese antecedente el que ejerce la función de S, como en (8 a-b), o de OD, como en (14 a-b).



12'. /sujeto pospuesto: cláusula relativa/

a. **quedó** muy sombría, *lo que era aquella habitación*

(Las Palmas H)

b. si **hubieran sido** el cincuenta por ciento *los que tomaban*

*esa decisión* (Santiago de Compostela H)

Las frecuencias absolutas de los sujetos que son FNs léxicas o cláusulas se presentan en el cuadro 3:

	FN				Cláusula				Total
	Determinada		Indeterminada		Sustantiva		Relativa		
	Ant.	Pos.	Ant.	Pos.	Ant.	Pos.	Ant.	Pos.	
Alcalá H	11	1	10	2	0	0	0	0	24
Alcalá M	32	15	0	1	0	0	0	0	48
Buenos Aires H	12	3	3	1	0	0	2	0	21
Buenos Aires M	30	7	4	2	0	0	2	1	46
Caracas H	14	7	3	3	0	0	0	0	27
Caracas M	13	4	3	0	0	0	1	0	21
Granada H	11	6	2	1	0	0	0	0	20
Granada M	13	12	3	3	0	1	0	0	32
La Habana H	16	15	1	5	2	0	0	0	39
La Habana M	35	10	0	5	1	0	0	0	51
Las Palmas H	33	9	0	2	0	0	0	1	45
Las Palmas M	35	3	0	2	0	2	0	0	42
Lima H	20	4	0	3	0	0	0	0	27
Lima M	23	14	1	4	1	4	0	0	47
Madrid H	16	3	0	2	0	0	2	0	23
Madrid M	16	13	0	4	0	0	0	0	33
México H	17	4	0	7	0	1	0	0	29
México M	18	14	0	6	0	0	0	0	38
Santiago Chile H	23	12	0	1	0	0	0	1	37
Santiago Chile M	6	2	0	0	2	0	0	0	10
Santiago Com. H	3	7	0	3	3	0	3	2	21
Santiago Com. M	8	3	1	2	0	1	1	0	16
Total	405	168	31	59	9	9	11	5	697

Cuadro 3. Uso de los sujetos FN y clausular

En el cuadro 3 puede verse que, de las FNs determinadas, 405 son antepuestas (71%) y 168 pospuestas (29%). Los porcentajes se invierten cuando se trata de FNs indeterminadas puesto que 31 aparecen antepuestas (34%) y 59, pospuestas (66%).

Los resultados sobre las FNs sujeto son más o menos esperables: si prototípicamente los sujetos suelen aparecer en posición preverbal y suelen representar entidades conocidas por los interlocutores, es natural que las FNs determinadas aparezcan más frecuentemente en esa posición que en posición postverbal, sobre todo si funcionan como sujetos de una oración copulativa (ver ej. 9a-b). Los casos de FNs determinadas y pospuestas se dan sobre todo con los verbos de existencia y de apreciación psicológica, como puede verse en (9'a-b) (Contreras 1976, Bentivoglio 1988). En cuanto a las FNs indeterminadas, todo hace pensar que su considerable porcentaje de posposición (66%) se debe a que su carácter indeterminado las hace entrar en competencia con los ODs que, como veremos más adelante, suelen ser indeterminados y aparecen a la derecha del verbo (Morales 1983: 477, *inter alia*)<sup>19</sup>.

Terminamos observando que hay relativamente pocos sujetos clausulares (ej. 11-12'), tanto en lo que corresponde a una cláusula sustantiva (9 de anteposición, 50%, y 9 de posposición, 50%) como a una cláusula relativa (11 de anteposición, 69%, y 5 de posposición, 31%).

## 4.2. El objeto directo

### 4.2.1. Los pronombres

Todos los ODs pronominales, en cualquiera de sus formas, han sido separados en tres grupos: el primero contiene los pronombres personales átonos (ej. 13 y 13'), el segundo corresponde a los pronombres relativos (14), y el tercero incluye los demás pronombres (ej. 15 y 15'):

- 13./OD antepuesto: pronombre personal átono (clítico)/
- a. él *me veía* y me decía cosas (Caracas M)
  - b. ya *lo entiende* (México M)

<sup>19</sup> Creemos que una de las razones más importantes para explicar la semejanza entre los Ss y los ODs nominales indeterminados se debe a lo que Du Bois (1987 y 2003) cree que es la base discursiva de la estructura argumental preferida. Según su hipótesis, todas las lenguas son ergativas, aun cuando no manifiesten morfológicamente los rasgos de la ergatividad. En efecto, y nos limitamos aquí a dos lenguas románicas, tanto Dutra (1987) como Bentivoglio (1992), al aplicar la hipótesis de Du Bois al portugués de Brasil y al español de Caracas, respectivamente, han encontrado en esas lenguas varios rasgos sintáctico-semántico-pragmáticos, aunque no morfológicos, propios de la ergatividad, que relacionan los ODs con los Ss de las oraciones intransitivas.

- 13'. /OD pospuesto: pronombre personal átono (clítico)/  
 a. **verlos** comer me encanta (La Habana H)  
 b. **tienen que aprenderlo** en, en la escuela (Madrid M)
14. /OD antepuesto: pronombre relativo/  
 a. a la edad *que* **tengo** (Buenos Aires H)  
 b. han sido más los problemas *que* te **ocasiona** el no estar ahí (México M)
15. /OD antepuesto: otros pronombres /  
 a. *eso* no **podría** ... no, no me imagino (Santiago, Chile M)  
 b. no sé si *algo* **tenemos** los gallegos (Santiago Com. M)
- 15'. /OD pospuesto: otros pronombres/  
 a. así como tenés el Photoshop **tenés otros** (Buenos Aires M)  
 b. nadie m(e) **ha dicho na(da)** (Granada M)

En el cuadro 4 se ofrecen los resultados correspondientes a los ODs pronominales:

	Pronombre personal átono		Pronombre relativo	Otros pronombres		Total
	Ant.	Pos.	Ant.	Ant.	Pos.	
Alcalá H	32	1	7	0	0	40
Alcalá M	13	4	4	0	3	24
Buenos Aires H	19	1	17	0	4	41
Buenos Aires M	15	1	5	0	2	23
Caracas H	7	2	2	0	1	12
Caracas M	16	5	1	0	2	24
Granada H	20	5	3	0	1	29
Granada M	29	2	4	0	11	46
La Habana H	25	9	7	0	0	41
La Habana M	22	1	4	0	0	27
Las Palmas H	18	3	4	0	1	26
Las Palmas M	25	2	2	0	4	33
Lima H	9	1	10	0	2	22
Lima M	17	1	4	0	2	24
Madrid H	19	2	6	0	0	27
Madrid M	22	6	2	0	4	34
México H	15	1	4	0	2	22
México M	7	1	3	0	4	15
Santiago Chile H	16	0	3	0	1	20
Santiago Chile M	22	0	0	1	1	24
Santiago Com. H	10	4	7	0	5	26
Santiago Com. M	23	2	5	1	0	31
Total	401	54	104	2	50	611

Cuadro 4. Uso del OD pronominal

No hay en el corpus analizado ningún pronombre personal tónico. Los pronombres con función de OD son todos átonos (clíticos) y, cuando acompañan a un solo verbo, su posición antepuesta o pospuesta está gramaticalmente condicionada por el verbo: si el verbo está conjugado, el clítico correspondiente al OD aparece antepuesto (ejs. 13a-b) y, si el verbo no está conjugado, el OD aparece pospuesto (ejs. 13'a-b).

Los resultados del cuadro 4 indican un considerable empleo de los OD pronominales (611 casos). Su variabilidad posicional con respecto al verbo no es relevante con los relativos pues está condicionada por la sintaxis del español. Tampoco lo es con los pronombres átonos por lo que se mencionó en el párrafo anterior. La única excepción se da cuando el clítico acompaña a una perífrasis verbal o a un determinado conjunto de verbos; en esos casos, es teóricamente posible tanto la anteposición del clítico como su posposición (NGRALE §16.12). Seguidamente se ilustran ambas posibilidades: en (16a-b) el clítico está antepuesto y en (16c-d), pospuesto.

16. /Pronombre clítico con perífrasis verbal o conjunto de verbos/
  - a. y esa casa se *la fueron haciendo* (Granada H)
  - b. y *las estoy viendo* por televisión (Madrid M)
  - c. pero bueno, **tratamos de sobrellevarlo** lo mejor posible (La Habana H)
  - d. **seguimos queriéndonos** con el mismo ... (La Habana H)

En el corpus se han encontrado 41 casos de este tipo, de los cuales hay 31 clíticos antepuestos (78%), y 9 pospuestos (22%). Estos resultados concuerdan con los de Zabalegui (2008), quien, en su estudio sobre la variación posicional de los clíticos en el español de Caracas, encuentra que, en la lengua hablada, la anteposición de los clíticos alcanza el 80% de los casos.

En el cuadro 4 se puede notar que la posición de los pronombres relativos es obligatoriamente antepuesta. En cuanto a los “otros pronombres”, solo hay dos casos de anteposición (4%), *vs.* 50 de posposición (96%), lo cual ha de interpretarse como que los pronombres pertenecientes a esa categoría que funcionan como OD tienen una fuerte tendencia a la posposición. Esta tendencia seguramente está motivada por el hecho de que la mayoría de ellos son cuantitativos indeterminados (ejs. 15' a-b). Sobre la tendencia a colocar los pronombres cuantitativos en posición posverbal, vale la pena señalar que esto no solo sucede con los ODs; recuérdese que, en el cuadro 2, el porcentaje de sujetos pronominales cuantitativos pospuestos era superior al de cualquier otro tipo de sujeto pronominal.

Antes de concluir este apartado vale la pena comentar que, dado el elevado uso del sujeto pronominal *yo* (ver cuadro 2), cabe preguntarse si verdaderamente ese uso corresponde al egocentrismo de la lengua hablada (*yo quiero, yo creo, yo hago*). Una manera de comprobarlo es observar qué sucede cuando se trata del OD. Puesto que esta función gramatical está semánticamente asociada al paciente, al que recibe la acción del sujeto, un uso elevado de *me* contrarrestaría la tendencia observada a emplear la 1ª persona del singular como sujeto. Los resultados indican que, de los 455 (antepuestos y pospuestos) pronombres personales que funcionan como OD, tan solo hay 57 casos de *me* (13%). Esta cantidad, muy inferior a los 395 casos (antepuestos y pospuestos) casos de *yo*, nos permite confirmar que el hablante suele referirse a sí mismo mucho más cuando es agente que cuando es paciente.

#### 4.2.2. Las frases nominales léxicas y las cláusulas

Las frases nominales léxicas con función de OD pueden ser determinadas (ejs. 17-17') o indeterminadas (ejs. 18-18'); los ODs clausulares pueden ser cláusulas sustantivas (ejs. 19-19') o relativas (ejs. 20-20'):

17. /OD antepuesto: FN determinada/  
(no se han encontrado ejemplos)
- 17'. /OD pospuesto: FN determinada/
  - a. **hicimos** la ruta patriótica también (Caracas H)
  - b. yo **experimenté** la crisis pues de los, de los ochenta (Lima H)
18. /OD antepuesto: FN indeterminada/
  - a. y no sé, **algún emprendimiento haría** (Buenos Aires M)
  - b. (no se han encontrado ejemplos)
- 18'. /OD pospuesto: FN indeterminada/
  - a. **empecé a a tener** trabajos particulares (Buenos Aires H)
  - b. **hemos cambiado** una persiana (Madrid H)
19. /OD antepuesto: cláusula sustantiva/  
(no se han encontrado ejemplos)
- 19'. /OD pospuesto: cláusula sustantiva/
  - a. yo **veo** que cada vez ve- vive más gente (Alcalá M)
  - b. yo le **decía** a mi mamá: "Es que necesito para un libro" (México H)

20. /OD antepuesto: cláusula relativa/  
(no se han encontrado ejemplos)

20'. /OD pospuesto: cláusula relativa/

a. probablemente te **recuerden** *lo que es la Rúa Nova de Abaixo* (Santiago de Com H)

b. **ver** un poco *lo que ha pasado el día anterior* (Alcalá H)

Las frecuencias absolutas de los ODs que son FNs o bien cláusulas se presentan en el cuadro 5:

	FN				Cláusula				Total
	Determinada		Indeterminada		Sustantiva		Relativa		
	Ant.	Post.	Ant.	Post.	Ant.	Pos.	Ant.	Post.	
Alcalá H	0	15	0	12	0	31	0	2	60
Alcalá M	0	21	0	18	0	2	0	0	41
Buenos Aires H	0	18	0	34	0	12	0	1	65
Buenos Aires M	0	19	1	33	0	16	0	2	71
Caracas H	0	16	0	41	0	5	0	0	62
Caracas M	0	35	0	39	0	12	0	3	89
Granada H	0	16	0	23	0	23	0	2	64
Granada M	0	21	0	31	0	20	0	3	75
La Habana H	0	13	0	11	0	5	0	0	29
La Habana M	0	15	0	30	0	6	0	1	52
Las Palmas H	0	18	0	19	0	7	0	1	45
Las Palmas M	0	22	0	41	0	11	0	0	74
Lima H	0	18	0	20	0	17	0	0	55
Lima M	0	27	0	27	0	5	0	0	59
Madrid H	0	14	0	30	0	20	0	0	64
Madrid M	0	18	0	16	0	3	0	0	37
México H	0	12	0	24	0	16	0	0	52
México M	0	8	0	22	0	11	0	1	42
Santiago Chile H	0	11	0	29	0	22	0	1	63
Santiago Chile M	0	15	0	30	0	16	0	0	61
Santiago Com. H	0	9	0	35	0	18	0	4	66
Santiago Com. M	0	14	0	37	0	15	0	1	67
Total	0	375	1	602	0	293	0	22	1293

Cuadro 5. Uso del OD nominal o clausular

Los datos del cuadro 5 indican que los ODs incluidos están casi siempre en posición posverbal. Dentro de ese grupo, sin embargo, los 602 casos de posposición con una FN indeterminada (62%) son bastante más numerosos que los 375 con una FN determinada (38%). La mayor frecuencia de los ODs posverbales se justifica precisamente por su carácter indeterminado: estos constituyentes se colocan frecuentemente a la derecha del verbo porque podrían o bien transmitir información desconocida para el interlocutor o bien constituir una colocación<sup>20</sup> del tipo *dar hambre*, *dar sueño*, etc. (ej., “no me pueden **dar plaza** hasta que se cumpla **la segunda parte de** la oposición”, Granada M).

Con respecto a la anteposición nominal del OD, cabe señalar que solo hemos registrado un caso en el corpus, el ilustrado en (18a), con una FN indeterminada. La falta de otros ODs antepuestos se debe a que, como puede verse en la sección 3, dedicada a la Metodología, en las cláusulas de dislocación hemos considerado que el OD corresponde al clítico situado en el interior de la cláusula y no al elemento dislocado, que consideramos extraclausular<sup>21</sup>. Sobre el tema de la dislocación, sin embargo, creemos conveniente informar, de manera adicional, que en el corpus encontramos 29 casos de dislocación a la izquierda del OD y ninguno de dislocación a la derecha. Esos casos se distribuyen de la manera siguiente: hay 19 casos en los que el constituyente dislocado es un FN determinada (21a-b), 5 casos de pronombre demostrativo (21c-d), y 2 casos de pronombre indeterminado<sup>22</sup>; en los ejemplos que siguen, el constituyente dislocado aparece subrayado y el clítico que lo representa, en cursiva:

21. /OD dislocados a la izquierda/  
 a. la tapia esa *la* **tendría que poner** en otro sitio (Granada M)  
 b. el frijol chino *lo* **mantienes** aparte (La Habana H)  
 c. ya eso *lo* **tenía** incorporado (Buenos Aires H)

<sup>20</sup> Las colocaciones son conjuntos de palabras cuya unión está determinada por el uso, si bien su significado es más o menos predecible del significado de las partes (Sedano 2011: 71).

<sup>21</sup> Algunos autores como Morales (1983) y Bogard (2010) consideran que, en las construcciones dislocadas de OD, esta función la desempeña no ya el clítico correspondiente situado en el interior de la construcción, sino la FN correferencial con dicho clítico, que puede aparecer antepuesta (Dislocación a la izquierda) o pospuesta (Dislocación a la derecha).

<sup>22</sup> El único caso de dislocación de una FN indeterminada es el siguiente: “me imagino que fríos intensos *los* ha habido siempre” (Alcalá H). No incluimos este caso en el corpus porque el verbo de la construcción es *haber*, que fue excluido del análisis por lo señalado en el punto viii del apartado 3.



- d. eso **tiene que saberlo** ella (Caracas M)
- e. todo lo **hacíamos** en carro (Caracas H)
- f. ya algunos **los tiene** puesto nuevos (Granada M)

Volviendo al cuadro 5, puede observarse que los ODs clausulares están siempre pospuestos al verbo y que las cláusulas sustantivas son mucho más frecuentes que las relativas sin antecedente: 293 casos = 93% *vs.* 22 casos = 7%, respectivamente.

### 4.3. Comparación de los resultados

En esta sección ofrecemos tres cuadros, uno que recoge en forma sintética todas las posibilidades de uso del sujeto en cualquiera de sus formas, otro que hace otro tanto con el OD, y un tercero en el que se comparan los resultados de los primeros dos cuadros. En algunas columnas de los cuadros que siguen, el número de casos es tan reducido que en el total aparece 0%, lo cual ha de interpretarse como que el porcentaje es inferior al 1%.

#### 4.3.1. Sujetos

El cuadro 6 ofrece los resultados relacionados con el uso de todas las formas que cumplen el rol de sujeto en el corpus analizado:

Ciudades	Ø		se		Pronombre				FN				Cláusula sustantiva o relativa				TOTAL	
		%	Ant.	%	Ant. <sup>1</sup>	%	Pos.	%	Ant.	%	Pos.	%	Ant.	%	Pos.	%		%
Alcalá H	198	79%	0	0%	27	11%	2	1%	21	8%	3	1%	0	0%	0	0%	251	4,91%
Alcalá M	134	62%	0	0%	32	15%	2	1%	32	15%	16	7%	0	0%	0	0%	216	4,23%
B. Aires H	148	66%	0	0%	51	23%	3	1%	15	7%	4	2%	2	1%	0	0%	223	4,36%
B. Aires M	152	61%	0	0%	51	20%	0	0%	34	14%	9	4%	2	1%	1	0%	249	4,87%
Caracas H	169	72%	0	0%	35	15%	3	1%	17	7%	10	4%	0	0%	0	0%	234	4,58%
Caracas M	144	59%	0	0%	80	33%	0	0%	16	7%	4	2%	1	0%	0	0%	245	4,79%
Granada H	157	63%	3	1%	67	27%	2	1%	13	5%	7	3%	0	0%	0	0%	249	4,87%
Granada M	163	69%	0	0%	39	17%	2	1%	16	7%	15	6%	0	0%	1	0%	236	4,62%
La Habana H	152	63%	0	0%	47	20%	3	1%	17	7%	20	8%	2	1%	0	0%	241	4,72%
La Habana M	167	64%	0	0%	41	16%	1	0%	35	13%	15	6%	1	0%	0	0%	260	5,09%
Las Palmas H	133	56%	0	0%	58	24%	2	1%	33	14%	11	5%	0	0%	1	0%	238	4,66%
Las Palmas M	163	63%	0	0%	53	20%	2	1%	35	13%	5	2%	0	0%	2	1%	260	5,09%
Lima H	156	67%	1	0%	48	21%	1	0%	20	9%	7	3%	0	0%	0	0%	233	4,56%
Lima M	163	62%	1	0%	52	20%	0	0%	24	9%	18	7%	1	0%	4	2%	263	5,15%
Madrid H	155	69%	2	1%	41	18%	3	1%	16	7%	5	2%	2	1%	0	0%	224	4,38%
Madrid M	161	72%	0	0%	28	13%	1	0%	16	7%	17	8%	0	0%	0	0%	223	4,36%
México H	140	60%	0	0%	57	25%	6	3%	17	7%	11	5%	0	0%	1	0%	232	4,54%
México M	130	66%	0	0%	26	13%	2	1%	18	9%	20	10%	0	0%	0	0%	196	3,84%
Sant., Chile H	140	60%	1	0%	51	22%	3	1%	23	10%	13	6%	0	0%	1	0%	232	4,54%
Sant., Chile M	184	85%	1	0%	20	9%	2	1%	6	3%	2	1%	2	1%	0	0%	217	4,25%
Sant. Com. H	118	65%	0	0%	41	23%	1	1%	3	2%	10	6%	6	3%	2	1%	181	3,54%
Sant. Com. M	137	66%	0	0%	53	26%	1	0%	9	4%	5	2%	1	0%	1	0%	207	4,05%
Total	3364	66%	9	0%	998	20%	42	1%	436	9%	227	4%	20	0%	14	0%	5110	100%

Cuadro 6. Resultados generales sobre el sujeto

Los resultados del cuadro 6 indican que los sujetos implícitos son los más numerosos del corpus ya que alcanzan el 66% de los sujetos de la muestra<sup>23</sup>. Siguen en orden de frecuencia los sujetos pronominales, por lo general antepuestos (20%)<sup>24</sup>; con excepción del clítico *se*, todos los demás pronombres encontrados son tónicos; dentro de esos pronombres, recordemos que juega un papel importante el de 1ª persona *yo* (ver cuadro 2). Después de los pronombres, destacan porcentualmente las FNs; aunque buena parte de ellas están antepuestas (9%), el porcentaje de FNs pospuestas cubre el 4% del corpus. Los sujetos clausulares, tanto antepuestos (20 casos) como pospuestos (14 casos), son extremadamente reducidos: no llegan ni al 1% del corpus. La poca relevancia porcentual de los sujetos clausulares se debe seguramente a su poca compatibilidad con los sujetos prototípicos, así como a las características coloquiales de las entrevistas.

Dado que los sujetos prototípicos suelen ser entidades conocidas por el interlocutor, es natural que la mayoría de ellos sean implícitos o pronombres, categoría esta que, cuando se trata de personales y demostrativos, corresponde a entidades conocidas. El empleo de las FNs, aunque más reducido, podría interpretarse así: su anteposición se da sobre todo cuando éstas son entidades que el interlocutor conoce o puede conocer (FNs det.); su posposición puede darse por distintas razones, entre ellas, que representen entidades que aparecen por primera vez en el discurso (FNs indet.), o bien que, constituyendo información conocida (FN det.), entren en competencia (contrastiva o no) con otras unidades del discurso; en esos casos, el hablante desecha las formas  $\emptyset$  y pronominales a causa de su eventual ambigüedad, y adopta una FN más informativa para una adecuada identificación de la entidad sujeto.

#### 4.3.2. Objeto directo

En el cuadro 7 presentamos los resultados obtenidos acerca de la presencia de pronombres y frases nominales léxicas que desempeñan el papel gramatical de OD:

<sup>23</sup> En el corpus de Bogard (2010: 109), el 58% de los sujetos son implícitos.

<sup>24</sup> Cabe señalar, sin embargo, que de ese 20%, correspondiente a 998 casos, hay 264 que están obligatoriamente antepuestos porque corresponden a pronombres relativos (cuadro 2).

Ciudades	Pronombre				FN				Cláusula sustantiva o relativa				Total	
	Ant.	%	Post.	%	Ant.	%	Post.	%	Ant.	%	Post.	%		%
Alcalá H	39	39%	1	1%	0	0%	27	27%	0	0%	33	33%	100	5,20%
Alcalá M	17	26%	7	11%	0	0%	39	60%	0	0%	2	3%	65	3,40%
B. Aires H	36	34%	5	5%	0	0%	52	49%	0	0%	13	12%	106	5,60%
B. Aires M	20	21%	3	3%	1	1%	52	55%	0	0%	18	19%	94	5,00%
Caracas H	9	12%	3	4%	0	0%	57	77%	0	0%	5	7%	74	3,90%
Caracas M	17	15%	7	6%	0	0%	74	65%	0	0%	15	13%	113	5,90%
Granada H	23	25%	6	6%	0	0%	39	42%	0	0%	25	27%	93	4,90%
Granada M	33	27%	13	11%	0	0%	52	43%	0	0%	23	19%	121	6,40%
La Habana H	32	46%	9	13%	0	0%	24	34%	0	0%	5	7%	70	3,70%
La Habana M	26	33%	1	1%	0	0%	45	56%	0	0%	7	9%	79	4,20%
Las Palmas H	22	31%	4	6%	0	0%	37	52%	0	0%	8	11%	71	3,70%
Las Palmas M	27	25%	6	6%	0	0%	63	59%	0	0%	11	10%	107	5,60%
Lima H	19	25%	3	4%	0	0%	38	49%	0	0%	17	22%	77	4,00%
Lima M	21	25%	3	4%	0	0%	54	65%	0	0%	5	6%	83	4,40%
Madrid H	25	27%	2	2%	0	0%	44	48%	0	0%	20	22%	91	4,80%
Madrid M	24	34%	10	14%	0	0%	34	48%	0	0%	3	4%	71	3,70%
México H	19	26%	3	4%	0	0%	36	49%	0	0%	16	22%	74	3,90%
México M	10	18%	5	9%	0	0%	30	53%	0	0%	12	21%	57	3,00%
Sant. Chile H	19	23%	1	1%	0	0%	40	48%	0	0%	23	28%	83	4,40%
Sant. Chile M	23	27%	1	1%	0	0%	45	54%	0	0%	16	19%	85	4,40%
Sant. Com. H	17	18%	9	10%	0	0%	44	48%	0	0%	22	24%	92	4,80%
Sant. Com. M	29	30%	2	2%	0	0%	51	53%	0	0%	16	16%	98	5,10%
Total	507	27%	104	5%	1	0%	977	51%	0	0%	315	17%	1904	100%

Cuadro 7. Resultados generales sobre el uso del OD

Los resultados del cuadro 7 indican que los ODs más frecuentes corresponden a FNs pospuestas (977 = 51%). Siguen en orden de importancia los pronombres (antepuestos, 27%, y pospuestos, 5%) y, finalmente, los ODs clausulares pospuestos (17%). Al sumar los porcentajes obtenidos para cada tipo de OD, se puede constatar que la posposición alcanza el 73%, lo que debe interpretarse como que solamente uno de cada cuatro ODs precede al verbo y que, en la gran mayoría de estos casos, hay un fuerte condicionamiento sintáctico: el OD es un clítico y el verbo está conjugado.

#### 4.3.3. La comparación

La comparación de los resultados correspondientes a los cuadros 6 y 7 permite señalar que, si bien los sujetos y los ODs pueden estar representados por categorías similares, hay numerosas diferencias entre esas dos funciones sintácticas: i) los sujetos son mucho más frecuentes que los ODs (5110 casos de Ss = 73% vs. 1904 casos de ODs = 27%); esto es natural si se piensa que, mientras casi todas las cláusulas tienen sujeto (implícito o explícito), solo las cláusulas transitivas llevan OD; ii) la categoría de los pronombres se emplea con gran frecuencia para representar tanto a los sujetos como a los ODs; sin embargo, con excepción de los nueve casos del clítico *se*, todos los pronombres personales encontrados en el corpus que representan al sujeto son tónicos, en tanto que los pronombres personales que representan al OD son átonos; iii) las FNs, que pueden ser preverbales o posverbales cuando son sujeto, casi siempre están en posición posverbal cuando son OD; iv) los sujetos clausulares, aunque de escaso uso, pueden ser preverbales o posverbales, pero los ODs clausulares, bastante frecuentes, parecen estar siempre pospuestos.

#### 4.4. Resultados extralingüísticos

En la presente investigación se han analizado solo los resultados relativos a la ciudad y al sexo del hablante, ya que la edad y el nivel educacional son similares en todos los hablantes.

En el cuadro 1 se ha observado poca variación en la frecuencia de los sujetos implícitos en cada ciudad ¿Podría pensarse que esa variación, aunque pequeña, indica diferencias dialectales? Creemos que no: el hecho de que el porcentaje más elevado de sujetos implícitos se dé en el hablante hombre de Alcalá de Henares (5,89%) queda compensando por el menor porcentaje de la hablante mujer de esa misma ciudad (3,98%). La similitud de los resultados nos lleva a pensar que el mayor número de sujetos implícitos no parece estar condicionado por la ciudad sino seguramente por el tipo de discurso.

Por razones de espacio, no ofrecemos en esta sección un cuadro relativo a los sujetos explícitos (ver anexo 2). Sin embargo, sí hemos hecho cálculos para determinar si las ciudades donde se pronuncia la *-s* final de palabra tienen menos sujetos explícitos que las ciudades donde la *-s* se aspira o se elide, sin que hayamos observado diferencias importantes. Si bien es cierto que los dos hablantes de Madrid, por ejemplo, ciudad en la que se pronuncia la *-s* final, emplean menos los sujetos explícitos (58 casos el hombre y 53, la mujer) de lo que se observa en una ciudad como Las Palmas de Gran Canaria (84 casos el hombre y 92, la mujer) donde se aspira o se elide la *-s* final, también es cierto que en Caracas, ciudad donde hay aspiración, pero no total elisión de *-s*, los resultados son contradictorios (48 casos el hombre, pero 91 la mujer). Un dato interesante es que, en Santiago de Chile, el hablante hombre emplea 77 sujetos explícitos y la mujer, tan solo 26. La diferencia es tan notable que no cabe atribuir la diferencia entre sujetos explícitos e implícitos a usos dialectales, sino más bien al tema del discurso, al número de entidades de las que se habla y quizá también al idiolecto particular de cada hablante.

Serrano (2013) observa que, en su corpus de estudio, las mujeres tienden a posponer el sujeto más que los hombres; motivadas por esa observación, hemos estudiado ese factor después de haber eliminado de nuestros cálculos los sujetos explícitos obligatoriamente antepuestos (*se* y pronombres relativos). Los resultados son poco contundentes: si bien es cierto que las mujeres posponen las FNs sujeto y los sujetos clausulares más que los hombres (FNs: 126 casos de las mujeres *vs.* 101 de los hombres; sujeto clausular: 9 casos de las mujeres y 5 de los hombres), también es cierto que la cantidad de pronombres pospuestos es casi idéntica en los dos grupos: 9 casos en las mujeres y 10 en los hombres. No hay suficientes casos, según puede verse en el cuadro 2, como para llegar a una conclusión fidedigna.

Con respecto a los resultados del OD no se observan diferencias dialectales que merezcan ser consideradas.

## 5. CONCLUSIONES

La presente investigación arroja resultados que, aun cuando son previsibles, permiten confirmar los hallazgos de muchos otros estudios:

- i) A pesar de su flexibilidad posicional, los sujetos tienden a anteponerse al verbo y los ODs, a posponerse.

- ii) Los sujetos más numerosos son los implícitos.
- iii) El número de sujetos (implícitos y explícitos) duplica ampliamente el de ODs.
- iv) Los sujetos explícitos son en su gran mayoría preverbales.
- v) Aunque los sujetos clausulares son extremadamente reducidos, el porcentaje de casos de anteposición es proporcionalmente menos elevado que el de los otros tipos de sujetos.
- vi) El sujeto *yo* se emplea mucho más que cualquier otro sujeto que sea pronombre personal.
- vii) Con excepción del sujeto impersonal *se*, los pronombres personales que representan al sujeto son siempre tónicos, en tanto que los que representan al OD son átonos.
- viii) El OD, en su representación como pronombre personal átono, es preverbal con los verbos conjugados y posverbal con los verbos no conjugados.
- ix) Cuando el clítico que funciona como OD modifica a una perífrasis verbal o bien a un determinado conjunto de verbos, existe tendencia a la anteposición del clítico.
- x) Los ODs que son FNs o cláusulas aparecen casi categóricamente después del verbo.
- xi) No parecen darse diferencias dialectales ni tampoco relativas al género de los hablantes.

Antes de concluir, quisiéramos aclarar que aunque, en principio, los temas de las entrevistas estaban fijados por los organizadores del proyecto PRESEEA, a la hora de realizar las grabaciones y, por razones muy variadas, se pasaba fácilmente a otros temas. A ello se debe el que, mientras algunos hablantes cuentan sus experiencias vitales, sus recuerdos, o manifiestan sus creencias, otros narran eventos históricos, políticos, económicos o de otra índole. De ahí que algunas entrevistas abunden en el uso de la 1ª persona del singular, mientras que en otras el uso de esta persona disminuye y aumenta el de la 3ª persona, sobre todo el representado por las frases nominales. En cualquier caso, y a pesar de los diferentes temas tratados en las entrevistas, las numerosas semejanzas en los resultados dan una idea de la homogeneidad de nuestra lengua, no importa el continente o la ciudad donde ésta se habla.

## REFERENCES

- Aijón Oliva, Miguel Ángel y María José Serrano (2010): “El hablante en su discurso: expresión y omisión del sujeto de *creo*”, *Oralia* 13, pp. 7-38.
- Alfonso Vega, Milagros y Chantal Melis (2010): “La posición del sujeto en la oración intransitiva del español”, en Sergio Bogard (ed.), pp. 39-68.
- Balash, Sonia (2008): “La conectividad discursiva en el discurso interactivo”, en Joyce Bruhn y Elena Valenzuela (eds.): *Selected proceedings of the 10th Hispanic Linguistic Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 301-311.
- Barrenechea, Ana María y Alicia Alonso (1977): “Los pronombres personales sujetos en el español hablado en Buenos Aires”, en Juan Manuel Lope Blanch (ed.): *Estudios sobre el español hablado en las principales ciudades de América*. México, DF: Universidad Nacional Autónoma de México, pp. 333-349.
- Bentivoglio, Paola (1987): *Los sujetos pronominales de primera persona en el español de Caracas*. Caracas: Universidad Central de Venezuela.
- (1988): “La posición del sujeto en el español de Caracas: un análisis de los factores lingüísticos y extralingüísticos”, en Robert M. Hammond y Melvyn Resnick (eds.): *Studies in Caribbean Spanish Dialectology*. Washington, DC: Georgetown University Press, pp. 14-23.
  - (1989a): “Función y significado de la posposición del sujeto nominal en el español hablado”, en *Estudios sobre español de América y lingüística afroamericana*. Bogotá: Instituto Caro y Cuervo, pp. 40-58.



- (1989b): “La posición del sujeto en las cláusulas copulativas en el español de Caracas”, en *Actas del VII Congreso de la Asociación de Lingüística y Filología de la América Latina (ALFAL)*, tomo II. Santo Domingo, Rep. Dominicana: Asociación de Lingüística y Filología de la América Latina (ALFAL), Filial Dominicana, pp. 173-196.
  - (1992): “La estructura argumental preferida del español hablado”, en Amparo Morales y María Vaquero (eds.): *Estudios sobre el español actual. Homenaje a Humberto López Morales*. San Juan, PR: Universidad de Puerto Rico, pp. 107-120.
  - (1993): “Full NPs: A discourse profile”, en William J. Ashby, Marianne Mithun, Giorgio Perissinotto y Eduardo Raposo (eds.): *Linguistic perspectives on the Romance languages (Selected papers from the XXI Linguistic Symposium on Romance Languages; Santa Barbara, February 21-24, 1991)*. Amsterdam/Philadelphia: John Benjamins, pp. 211-224.
  - (2003): “Orden de palabras en español: un análisis sintáctico-semántico-pragmático del sujeto”, *Lexis* XXVII, 1-2, pp. 235-260.
  - y Francesco D’Introno (1978): “Orden de palabras y posición del sujeto en el español de Caracas: un estudio sociolingüístico”, ponencia presentada ante el V Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina, Casa de Bello, Caracas, 5-9 de enero de 1978.
  - y Francesco D’Introno (1989): “Orden de palabras y posición del sujeto en el español de Caracas”, en Iraset Páez Urdaneta, Fernando Fernández y Luis Barrera Linares (comps.): *Estudios lingüísticos y filológicos en homenaje a María Teresa Rojas*. Caracas: Universidad Simón Bolívar, pp. 58-82.
  - y Elizabeth Weber (1986): “A functional approach to subject word order in spoken Spanish”, en Oswaldo Jaeggli y Carmen Silva-Corvalán (eds.): *Studies in Romance Linguistics*. Dordrecht: Foris Publications, pp. 23-40.
- Blanco Canales, Ana (1999): “Presencia/ausencia de sujeto pronominal de primera persona en español”, *Español Actual* 72, pp. 31-39.
- Bogard, Sergio (ed.) (2010): *Semántica, pragmática y prosodia: reflejos del orden de palabras en español*. México, DF: El Colegio de México.

- (2010): “La frase nominal de objeto directo antepuesta al verbo en español”, en Sergio Bogard (ed.), pp. 69-115.
- Bolinger, Dwight L. (1954-55): “Meaningful word order in Spanish”, *Boletín de Filología. Homenaje a Rodolfo Oroz* 7, pp. 45-56.
- Cameron, Richard (1992): *Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation*, tesis doctoral, University of Pennsylvania at Philadelphia.
- (1995): “The scope and limits of switch reference as a constraint on pronominal subject expression”, *Hispanic Linguistics* 6/7, pp. 1-27.
- y Nydia Flores Ferrán (2004): “Perseveration of subject expression across regional dialects of Spanish”, *Spanish in Context* 1, 1, pp. 41-65.  
<http://dx.doi.org/10.1075/sic.1.1.05cam>
- Cantero, Gustavo (1976): “Peculiaridades en el empleo del pronombre personal *yo* en el habla culta de la Ciudad de México”, *Anuario de Lingüística* 14, pp. 233-237.
- Comrie, Bernard (1981): *Language universals and linguistic typology*. Chicago: University of Chicago Press. Versión en español, *Universales del lenguaje y tipología lingüística*. Madrid: Gredos, 1988.
- Contreras, Heles. (1976): *A theory of word order with special reference to Spanish*. Amsterdam: North-Holland Publishing Co. Versión en español, *El orden de palabras en español*. Madrid: Cátedra, 1978.
- Croft, William (1990): *Typology and universals*. Cambridge, MA: Cambridge University Press.
- Davidson, Brad (1996): “Pragmatic weight and Spanish subject pronouns: The pragmatic and discourse uses of ‘tú’ and ‘yo’ in spoken Spanish”, *Journal of Pragmatics* 26, 4, pp. 543-565.  
[http://dx.doi.org/10.1016/0378-2166\(95\)00063-1](http://dx.doi.org/10.1016/0378-2166(95)00063-1)
- Delbecque, Nicole (1979): *Étude quantitative d’un phénomène de variation syntaxique: la position de sujets en espagnol*, tesis doctoral, Katholieke Universit at Leuven.
- (1987): *Problèmes et méthodes de l’étude de la variation syntaxique: le cas de la position du sujet en espagnol*. Leuven: Leuven University Press.
- (1991): *El orden de los sintagmas. La posición del regente*. Salamanca: Universidad de Salamanca.

- (2005): “El análisis de corpus al servicio de la gramática cognoscitiva: hacia una interpretación de la alternancia lineal SV / VS”. En Gabrielle Knauer y Valeriano Bellosta von Colbe (eds.): *Variación sintáctica en español: un reto para las teorías de la sintaxis*. Tübingen: Niemeyer, pp. 51-74.
- Dryer, Matthew S. (1992): “The greenbergian word order correlations”, *Language* 68, pp. 81-138.  
<http://dx.doi.org/10.2307/416370>  
<http://dx.doi.org/10.1353/lan.1992.0028>
- Du Bois, John W. (1987): “The discourse basis of ergativity”. *Language* 63, pp. 805-855.  
<http://dx.doi.org/10.2307/415719>
- (2003): “Argument structure: grammar in use”, en John W. Du Bois, Lorraine E. Kumpf y William J. Ashby (eds.): *Preferred Argument Structure: Grammar as architecture for function*. Amsterdam: John Benjamins, pp. 11-60.
- Dumont, Jenny (2006): “Full NPs as subjects”, en Nuria Segarra y Jacqueline Almeida (eds.): *Selected proceedings of the 9<sup>th</sup> Hispanic Linguistic Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 286-296.
- Dutra, Rosalia (1987): “The hybrid S-category in Brazilian Portuguese: Some implications for word order”, *Studies in Language* 11, 1, pp. 163-180.  
<http://dx.doi.org/10.1075/sl.11.1.07dut>
- Enríquez, Emilia V. (1984): *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
- (1986): “La presencia de los pronombres personales sujeto en el mundo hispánico: un estudio comparativo”, *Anuario de Lingüística Hispánica* XXIV, pp. 47-70.
- Fernández Soriano, Olga (1999): “El pronombre personal. Formas y distribuciones. Pronombres átonos y tónicos”, en Ignacio Bosque y Violeta Demonte (eds.): *Gramática descriptiva de la lengua española*, Vol. 1. Madrid: Espasa Calpe, pp. 1161-1266.
- Gili Gaya, Samuel (1961): *Curso superior de sintaxis española*. Barcelona: Vox.
- Greenberg, Joseph H. (1963): *Universals of language*. Cambridge, MA: The M.I.T. Press.

- Hatcher, Anna G. (1956): "Theme and underlying question: Two studies of Spanish word order", *Word* 12, Supplement # 3.
- Hidalgo Downing, Raquel (2003): *La tematización en español*. Madrid: Gredos.
- Hurtado, Luz Mercedes (2001): *La variable expresión del sujeto en el español de los colombianos y colombo-americanos residentes en el condado de Miami-Dade*, tesis doctoral, University of Florida.
- Lambrecht, Knud. (2001): "Dislocation", en Martin Haspelmath, Ekkehard König, Wulf Oesterreicher y Wolfgang Raible (eds.): *Language typology and language universals. An international handbook*, vol. 2. Berlin/ New York: Walter de Gruyter, pp. 1050-1078.
- Lu, Hui-Chuan (1997): "La ausencia y presencia del pronombre personal", *EPOS XIII*, pp. 117-133.
- Martín Butragueño, Pedro (2010): "La posición extrapredicativa de tema en la lengua hablada", en Sergio Bogard (ed.), pp. 117-183.
- Martínez Lara, José Alejandro (2012): *Estudio del sujeto variable de 3ª persona en el español de Caracas 2004-2010*, tesis de Maestría en Lingüística, Universidad Central de Venezuela, Caracas.
- Miró Vera, Ramona y Miguel Ángel de Pineda (1990): "Determinación sociolingüística de la presencia/ausencia del pronombre personal sujeto", en María Teresa Palet Plaja (ed.): *Sociolingüística andaluza 5: habla de Sevilla y hablas americanas*. Sevilla: Universidad de Sevilla, pp. 37-44.
- Morales, Amparo (1982): "La perspectiva dinámica oracional en el español de Puerto Rico", en Orlando Alba (ed.): *El español del Caribe*. Santiago, Rep. Dominicana: Universidad Católica Madre y Maestra, pp. 203-219.
- (1983): "Inversiones de objeto en el español de Puerto Rico", en *Philologica hispaniense: in honorem Manuel Alvar*, vol. 1. Madrid, Gredos, pp. 471-484.
  - (1986): "Algunos aspectos de gramáticas en contacto: la expresión del sujeto en el español de Puerto Rico", *Anuario de Letras* 24, pp. 71-85.
  - (1987): "La manifestación del sujeto en español: tendencias universales y variación dialectal", en Robert Hammond y Marguerite MacDonald (eds.): *Linguistic studies in honor of Bohdan Saciuk*. West Lafayette, IN: Learning Systems, Inc., pp. 87-103.

- Moreno Fernández, Francisco (1997): “Metodología del ‘Proyecto para el Estudio Sociolingüístico del Español de España y de Hispanoamérica’”, en Francisco Moreno Fernández (ed.), *Trabajos de Sociolingüística hispánica*. Alcalá de Henares: Universidad de Alcalá, pp. 137-161.
- (2005): “Corpus para el estudio del español en su variación geográfica y social”, *Oralia* 8, pp. 123-139.
- Ocampo, Francisco (2005): “El orden de palabras en español hablado: construcciones con un verbo cópula, un sintagma nominal sujeto y un sintagma nominal complemento”, en *Actas del XIV Congreso Internacional de la ALFAL (octubre de 2005)*. Monterrey: Universidad Autónoma de Nuevo León (CD).
- (2009): “El orden de palabras en el español hablado. La construcción Sujeto Verbo Objeto Directo”, en Montserrat Veyrat Rigat y Enrique Serra Alegre (eds.): *La lingüística como reto epistemológico y como acción social. Estudios dedicados al Profesor Ángel López García con ocasión de su sexagésimo aniversario*, vol. 1. Madrid: Arco Libros, pp. 501-511.
  - (en prensa): “No, pero siguen las supersticiones. Sí, las supersticiones siguen. El orden de palabras en rioplatense: factores cognitivos, sintácticos y pragmáticos”, en Laura Colantoni y Celeste Rodríguez Louro (eds.): *Perspectivas teóricas y experimentales sobre el español de la Argentina*. Frankfurt y Madrid: Iberoamericana / Vervuert.
- Pérez Brabandere, Vanessa (2010): *Los sujetos pronominales de 1ª persona en el español de Caracas 2004-20010*, tesis de Maestría en Lingüística, Universidad Central de Venezuela, Caracas.
- Posio, Pekka (2012): “Spanish subject pronoun usage and verb semantics revisited: First and second person singular subject pronouns and focusing of attention in spoken Peninsular Spanish”, *Journal of Pragmatics* 43, pp. 777–798.  
<http://dx.doi.org/10.1016/j.pragma.2010.10.012>
- Quijada, Milagros (2006): *Ø cantaba ... ¿él, ella o yo? Sujetos de 3ª persona en el habla de Caracas*, tesis de Maestría en Lingüística, Universidad Central de Venezuela, Caracas.
- Real Academia Española (1973): *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.

- y Asociación de Academias de la lengua española (2009): *Nueva gramática de la lengua española* (NGRALE). Madrid: Espasa Libros.
- Samper Padilla, José Antonio; Hernández Cabrera, Clara Eugenia; Troya Déniz, Magnolia (2006): “Presencia/ausencia del sujeto pronominal de primera persona singular en la norma culta de España”, en Concepción Company Company (ed.): *El español en América. Diatopía, diacronía e historiografía*. México, DF: Universidad Nacional Autónoma de México, pp. 87-109.
- Sedano, Mercedes (2011): *Manual de gramática del español, con especial referencia al español de Venezuela*. Caracas: Universidad Central de Venezuela.
- (2012): “Dislocación a la izquierda y a la derecha: semejanzas y diferencias”, en Tomás Jiménez Juliá, Belén López Meirama, Victoria Vázquez Rozas y Alexandre Veiga (eds.): *Cum corde et in nova grammatica. Estudios ofrecidos a Guillermo Rojo*. Santiago de Compostela: Universidade de Santiago de Compostela, pp. 791-805.
- Serrano, María José (2013): “Variación sociosituacional de la colocación del sujeto pronominal en textos conversacionales”, *Spanish in Context* 10,2, pp. 261-283.  
<http://dx.doi.org/10.1075/sic.10.2.04ser>
- Silva-Corvalán, Carmen (1977): *A discourse study of some aspects of word order in the Spanish spoken by Mexican-Americans in West Los Angeles*, M. A. thesis, University of California Los Angeles.
- (1982): “Subject variation in spoken Mexican-American Spanish”, en Jon Amastae y L. Elías Olivares (eds.): *Spanish in the United States: Sociolinguistic aspects*. New York: Cambridge University Press, pp. 93-120.
- (1983): “On the interaction of word order and intonation: Some OV constructions in Spanish”, en Flora Klein-Andreu (ed.): *Discourse perspectives on syntax*. New York: Academic Press, pp. 117-140.
- (1984): “Topicalización y pragmática en español”, *Revista Española de Lingüística* 14,1, pp. 1-19.
- (1997): “Variación sintáctica en el discurso oral: problemas metodológicos”, en Francisco Moreno Fernández (ed.), pp. 115-135.
- (2001): *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.

- (en prensa): “The acquisition of grammatical subjects by Spanish-English bilinguals”, en Ana M. Carvalho, Rafael Orozco y Naomi Lapidus Shin (eds.): *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Washington, DC: Georgetown University Press.
- Tomlin, Russell S. (1985): “Interaction of subject, theme, and agent”, en Jessica Wirth (ed.): *Beyond the sentence: Discourse and sentential form*. Ann Arbor, MI: Karoma, pp. 60-80.
- Travis, Catherine (2007): “Genre effects on subject expression in Spanish: Priming in narrative and conversation”, *Language Variation and Change* 19, 2, pp. 101-135.  
<http://dx.doi.org/10.1017/S0954394507070081>
- Vázquez Rozas, Victoria (2004): “Transitividad prototípica y uso”, *Boletín de Lingüística* 21, p. 92-115.
  - y José María García Miguel (2006): “Transitividad, subjetividad y frecuencia de uso”, en *Actes del VII Congrés de Lingüística General* (Barcelona, 18-21 de abril de 2006). Publ. CD-ROM (ISBN 84-475-2086-8).
- Zabalegui, Nerea (2008): “La posición de los pronombres átonos en construcciones con verbos no conjugados en el español actual de Caracas”, *Akados* 10,2, pp.87-107.



## **Anexos**

### **ANEXO 1**

Ciudad	Hombres	Mujeres
Alcalá de Henares	H32-031	M32-034
Buenos Aires	H32-Facundo	M32-Laura
Caracas	H32-085	M32-091
Granada	H32-007	M32-010
La Habana	H32-085	M32-091
Las Palmas de Gan Canaria	H32-009	M32-013
Lima	H32-051	M32-041
Madrid	H32-031	M32-034
Ciudad de México	H32-006	M32-055
Santiago de Chile	H32-085	M32-091
Santiago de Compostela	H32-003	M32-001

**Anexo 1.** Códigos de los hablantes



## ANEXO 2

	Antepuestos	%	Pospuestos	%	Total.
Alcalá H	38	88%	5	12%	43
Alcalá M	54	75%	18	25%	72
Buenos Aires H	58	89%	7	11%	65
Buenos Aires M	78	89%	10	11%	88
Caracas H	35	73%	13	27%	48
Caracas M	87	96%	4	4%	91
Granada H	70	89%	9	11%	79
Granada M	47	72%	18	28%	65
La Habana H	49	68%	23	32%	72
La Habana M	59	79%	16	21%	75
Las Palmas H	70	83%	14	17%	84
Las Palmas M	83	90%	9	10%	92
Lima H	56	87%	8	13%	64
Lima M	62	74%	22	26%	84
Madrid H	50	86%	8	14%	58
Madrid M	35	66%	18	34%	53
México H	68	79%	18	21%	86
México M	32	59%	22	41%	54
Santiago Chile H	60	78%	17	22%	77
Santiago Chile M	22	85%	4	15%	26
Santiago Com. H	25	66%	13	34%	38
Santiago Com. M	52	88%	12%	12%	59
Total	1190	81%	283	19%	1473

Anexo 2. Sujetos explícitos con flexibilidad posicional

## **El orden de palabras en cláusulas subordinadas relativas con un sujeto y un verbo transitivo**

Francisco Ocampo  
*University of Minnesota*

**Abstract:** This work presents a motivation for word order variation in relative clauses formed by a lexical NP subject and a transitive verb in spoken Rioplatense Spanish. The main correlating factor for the preverbal position of the subject is the topicality of the NP referent, which is a motivation also present in simple independent sentences. However, there are also other factors that correlate with the aforementioned word order in relatives. In the data all preverbal subjects appear in non-restrictive relative clauses, which also exhibit an independent tonal group. The explanation for this behavior would be that non-restrictive relative clauses have greater syntactic autonomy, which allows them to express a sentence topic that is different from the one expressed by the main clause. This syntactic factor correlates prosodically with the presence of an independent tonal group.

**Keywords:** Word order, relative clauses, spoken Rioplatense Spanish, topic, non-restrictive relatives, restrictive relatives.



## 1. INTRODUCCIÓN

En este trabajo se analiza el orden de palabras en cláusulas relativas formadas por un sujeto léxico<sup>1</sup> y un verbo transitivo, como se ejemplifica en (1) y (2)<sup>2</sup>:

(1) 16a8

M: pero arriba del aparador hay algo que ustedes

A: a:h () no: =

E: =una plancha sí. La plan:cha:

[

M: sí la plancha

A: ¿la plancha no?

M: la plancha (vieja)

A: y e ésto es lo que - no no la plancha () yo señalé eso es este: =

M: =/y la pla'ncha /**que Fra'ncisco le había traí'do**/ para para hacer o/ - y llega [...]

A: y ¿cómo vino a parar a tu ca:sa? [se dirige a L]

E: porque cuando ellas estaban sacando todas las cosas [...]

M: porque Amalia tenía que mudarse para que - para poder hacer lugar para las cosas de Amalia

[

E: estaba [...] yo me las llevo a casa, digo, yo las tengo acá y cuando los chicos [...] cuando los chicos se las quieren lleva:r=

M: =y Elena dijo mirá después () yo - mirá son de Francisco todas esas cosas, bueno, cuando Francisco las quiere si él le quiere cambiar el color () pero estaba toda oxidada la plancha que uno la toca y se resbala [...]

A: era de tu casa esto o era ¿de dónde e:ra? [se dirige a F] [...]

M: cuando desarmaron la ca:sa, Francisco la tra:jo, nena, cuando trajo tantas co:sas

<sup>1</sup> Es decir que no se incluyen sujetos pronominales.

<sup>2</sup> Las convenciones utilizadas en la transcripción aparecen al final del trabajo.

(2)1a15

A: este: son - que siempre tomo ahora ahora, son acidófilus.

M: [ ah sí,

que e - cla:ro, l:as pastillas para contra ácida.

(0.4)

A: no no::, son para los intestinos. /So'n los bichi^tos **que tie''nen los yogu'r.**/

En (1) el orden de los constituyentes de la cláusula relativa es SV, es decir que el sintagma nominal sujeto aparece en posición preverbal, mientras que en (2) el sintagma nominal sujeto aparece en posición postverbal, con un orden VS.

Las cláusulas subordinadas relativas han sido caracterizadas por Givón (1990:646) de la siguiente manera: “A relative clause codes a proposition one of whose participants is coreferential with the head noun that is modified by the clause”. Así, en (1) la cláusula relativa *que Francisco le había traído*, modifica al sintagma nominal *la plancha*, el cual a su vez está representado<sup>3</sup> por el pronombre relativo *que*, el cual cumple la función de objeto directo dentro de la cláusula relativa.

Tradicionalmente, las cláusulas relativas han sido divididas en dos clases: *especificativas* y *explicativas*. Esta distinción, introducida por los gramáticos de Port-Royal en su *Logique*, fue posteriormente reconocida por Bello (Brucart 1999:408). Las cláusulas relativas *especificativas*, también denominadas *restrictivas*, *determinantes* o *atributivas* y ejemplificadas en (2), “ejercen la función de complemento restrictivo del nombre o grupo nominal que actúa como antecedente” (Brucart 1999:410). Por su parte, las cláusulas relativas *explicativas*, llamadas también *incidentales* o *apositivas* y ejemplificadas en (1), no restringen la extensión del sintagma nominal sino que aportan información adicional sobre su referente (Brucart 1999:410).

Podría argumentarse, sin embargo, que en (1) *la plancha que Francisco le había traído* no se está ante una relativa explicativa sino una especificativa, es decir que la relativa no añade información adicional sobre el antecedente *la plancha* sino que su función es fundamentalmente restringir su significado, es decir identificar el referente de manera que se lo pueda distinguir de otros.

<sup>3</sup> Es decir que el sintagma nominal y el pronombre relativo son correferenciales.

Para justificar el análisis de la cláusula como explicativa hay que referirse con cierto detalle al contexto amplio en que se da la conversación. A y F están casados y viven en los EEUU. Han vuelto a La Plata, Argentina, de vacaciones y están visitando a una tía de A, la hablante E, acompañados de la madre de A, la hablante M. Antes de viajar a los EEUU y antes de casarse, F vivía con sus padres. Los padres de F se mudaron a una casa más chica y varias pertenencias familiares fueron llevadas por F a la casa de A, quien vivía con su madre M. Entre éstas estaba una antigua plancha a carbón, que F planeaba restaurar pero nunca llegó a hacerlo. Cuando A y F se trasladaron a los EEUU, una hermana de M, Amalia, se mudó a la casa de M. Para hacer lugar a las pertenencias de Amalia, varios objetos, entre ellos la plancha a carbón, pasaron a la casa de la otra hermana de M, la hablante E. Los participantes en la conversación conocen la historia de la plancha: nótese el empleo del artículo definido por los hablantes. La única salvedad es que A y F solamente ignoran la última parte de la historia: la plancha a carbón ahora está en la casa de E. Por lo tanto, la cláusula relativa de M no tiene la función de identificar al referente del sintagma nominal *la plancha*, sino que agrega información adicional sobre el origen de la plancha a carbón.<sup>4</sup>

Los primeros trabajos sistemáticos sobre el orden de palabras en castellano que utilizan corpus de datos reales, es decir no elicitados ni introspectivos (Ocampo 2010), son los de Kahane y Kahane (1950) y Hatcher (1956). Estos investigadores utilizan fuentes literarias. Además, Kahane y Kahane (1950) suplementan el corpus escrito con datos tomados de informantes. En este estadio primero, las generalizaciones a las que se llega son necesariamente provisionales.

Por otra parte, la primera lingüista que utiliza con exclusividad datos conversacionales del castellano es Silva-Corvalán (1977, 1983, 1984). Silva-Corvalán (1983) indica la relación entre el contorno tonal, la función pragmática de la oración y el orden de palabras. El trabajo pionero de Silva-Corvalán fue seguido por los artículos de Bentivoglio (1985) y Bentivoglio y Weber (1986). Estos estudios han sido los primeros en emplear factores sintácticos, pragmáticos y prosódicos para el análisis del orden de palabras en castellano. Los resultados de estas investigaciones, muy adelantadas para su tiempo, apuntan en la dirección correcta, ya que el análisis empírico de los datos de este artículo confirma las correlaciones mencionadas antes.

---

<sup>4</sup> Estas observaciones son posibles porque el autor es el hablante F, conoce a los otros hablantes, estuvo presente durante la conversación y participó en ella.

La presente investigación se basa en un corpus de 20 horas de conversaciones informales grabadas por el autor y Alicia Ocampo en 1986 y 1987, transcriptas siguiendo una notación conversacional (cf. Convenciones al final del trabajo). Los hablantes que participan son 36, de los cuales 20 son mujeres y 16 varones, con edades de entre 2 y 76 años. Los hablantes son todos nacidos y educados en La Plata, Argentina, miembros de una red social constituida por amigos o parientes de los investigadores. El autor, nativo también de La Plata, estuvo presente durante las grabaciones y participó en la conversación.

Para el análisis se consideraron todas las construcciones relativas que aparecieron en los datos, formadas exclusivamente por un verbo transitivo y un sintagma nominal sujeto, como las que se ejemplifican en (1) y (2). Toda relativa que contuviera elementos adicionales fue descartada. Para las construcciones analizadas también se anotó el grupo melódico (Cruttenden 1986; Sosa, 1999) y dentro de éste la prominencia prosódica con tres niveles, identificados como acento primario, secundario y terciario (cf. Convenciones). La prominencia prosódica fue registrada de forma impresionista, sin discriminar si es el resultado de intensidad, longitud, tono, o una combinación de estos factores. No se utilizó un análisis acústico pero los resultados son confiables. Hay que notar que el análisis acústico no identifica automáticamente sobre qué sílaba cae la prominencia principal, sino que despliega una serie de factores que el analista debe interpretar. Además, este tipo de análisis hay que hacerlo en conversaciones grabadas en una cámara insonorizada, sin ruido exterior. En los datos las conversaciones fueron grabadas en entornos naturales con ruidos que la mente humana puede filtrar pero la computadora no. En todo caso, si la prominencia prosódica comunica algo, debe poder ser percibida por el oyente. El hecho de estar analizando la propia variedad de español, estar familiarizado con el dialecto de los participantes y haber estado presente en las conversaciones, coloca al autor en la posición ideal para percibir la prominencia prosódica. Además, para asegurar la fidelidad de las notaciones prosódicas se presentó una muestra de las mismas a un hablante nativo rioplatense, quien confirmó los resultados.

Chafe (1994:63) distingue tres tipos de grupo melódico: *sustantivo*, *regulatorio* y *fragmentado*. Las unidades sustantivas aparecen como construcciones completas. Los grupos melódicos que aparecen en (1) y (2) son unidades sustantivas y son las que se tomaron en cuenta en el análisis. Las unidades regulatorias coinciden con partículas discursivas: *bueno*, *claro*,

*mirá vos*, etc. A causa de los criterios de selección (cláusulas relativas) estas unidades quedaron necesariamente fuera de las construcciones analizadas. Los grupos melódicos fragmentados son aquellos que por diversas razones el hablante o no los completa o los emite con interrupciones, como se ejemplifica en (3):

(3) 17a22

P: no, ahora l:o que yo te voy a decir, es de - un locutor **que tenía - canal nueve**. () Escandaloso al extremo. () Es de esos - es un hombre ¿cómo te puedo decir? grande como Carmelo nuestro () Corso Gómez

En este contexto es posible identificar una cláusula relativa, cuya forma sin autocorrecciones (Levinson 1983) sería la siguiente:

(4) un locutor que tenía canal nueve.

La construcción fue registrada para el análisis pero a causa de su carácter fragmentado no se intentó distinguir el grupo melódico. Hay en los datos un total de siete construcciones de este tipo donde no se estableció el grupo melódico.

## 2. PRESENTACIÓN DE LOS RESULTADOS

En los datos aparece un total de 42 cláusulas relativas. Estas cifras contrastan con la cantidad de oraciones simples independientes, con un sujeto y un verbo transitivo solamente, presentes en el corpus. Las frecuencias diferenciadas se muestran en el Cuadro 1:

Tipo de construcción	Cantidad	Frecuencia
Oraciones independientes	83/125	66.4
Clausulas relativas	42/125	33.6

**Cuadro 1.** Oraciones independientes y cláusulas relativas

El Cuadro 1 indica que las frecuencias de las construcciones relativas constituyen un tercio del total de construcciones con un verbo transitivo y un sujeto que aparecen como oraciones independientes o como cláusulas relativas, es decir 42/125, 33.6%.



La frecuencia de los dos órdenes de palabras, sujeto en posición preverbal y sujeto en posición postverbal, en las cláusulas relativas se indica en el Cuadro 2:

Orden de palabras	Cantidad	Frecuencia
Que S V	6/42	14.29
Que V S	36/42	85.71

**Cuadro 2.** Sujeto preverbal y sujeto postverbal en cláusulas relativas

El Cuadro 2 indica que la mayoría de las cláusulas relativas, 36/42, 85.71% presenta un sujeto postverbal y que solamente hay seis construcciones con un sujeto preverbal. En todas estas seis construcciones el referente del sintagma nominal sujeto en posición preverbal es el tópico de la cláusula, como puede observarse en (5):

(5) 9b9

P: y papá acostumbraba a sentarse en un sillón () en un sillón de hamaca de ésos de mimbre () y: fumaba en pipa () y el perro se sentaba así, al lado de él

[...]

L: pero cuando papá murió () este:, él se () dio cuenta ¿no? que faltó papá, entonces eh m recogía las zapatillas, las que: () /las últimas, /**que papá^ había usa'do**/, () agarraba eh la zapatilla, iba a buscar la zapatilla y se ponía abajo de alguna cama y se acostaba () sobre la zapatilla

El tópico de la conversación anterior es el referente del sintagma nominal *las zapatillas*. Por otra parte, el tópico de la cláusula relativa es el referente del sintagma nominal sujeto *papá*.

Al mismo tiempo, en todas las construcciones relativas con un sujeto en posición postverbal el referente del sintagma nominal sujeto presenta baja topicalidad, como puede observarse en (6) y (7):

(6) 2b8

N: lo que yo no quie:ro, cuando están juntos, cuando están juntos, no quiero que después digan huy los chicos fastidian. Los chicos molestan.

[

- P: deja:los. Pero vos dejalos, si ellos lo saben.  
[
- A: bueno, van a fastidiar de todas maneras. he he he [...]
- P: pero eso, no:, /e^so es el ro»l **que cumple el pa:'dre**/. El padre [...] el padre cumple el rol ese de retarlos. No, el abuelo no los puede retar. Porque los chicos se enoj
- (7) 19a4
- V: tonces m yo tirada en la cama de costado como me habían dicho e - lado dere e - izquierdo, diciéndole al flaco, (0.4) flaco, poné el camisón rosa.
- F: m hh
- V: tonces el flaco ponía el ca
- A: he
- V: /la's pantu»flas **que^ me regaló Patri"cia**/. [risas] viste, e:l - el champú: que está en el baño, el Sedal, qué sé yo, bueno, iba y venía - y así preparó to:do el bolso.

En el contexto previo a (6) el hablante P recrimina a N, su madre, que trata mal a sus nietos, los hijos de P; N responde indicando las razones de su proceder. El tópico de la cláusula relativa, coincide con el de la principal. Por su parte, el referente del sintagma nominal sujeto de la relativa, *el padre*, es de baja topicalidad y aparece en posición postverbal. En (7) la hablante V, que está próxima a dar a luz, cuenta a A y F un episodio sucedido unos días antes cuando tuvo contracciones y creyó equivocadamente que había comenzado el proceso de parto. En el trozo ejemplificado, V narra las instrucciones que en esa ocasión fue dando a su esposo para que le preparara la ropa para ir al hospital. Aquí el tópico de la cláusula subordinada es el referente del sintagma nominal, antecedente de la relativa, *las pantuflas*. El referente del sujeto de la relativa, *Patricia*, tiene baja topicalidad y aparece en posición postverbal. El ejemplo (2) donde el sujeto *los yogur* aparece en posición postverbal, también es de baja topicalidad; el tópico de la relativa *los bichitos* coincide con el de la cláusula principal.

Los datos presentados permiten afirmar que la motivación de la posición preverbal del sujeto del verbo transitivo en las cláusulas relativas es la topicalidad de su referente. Este factor también motiva la posición preverbal en las oraciones independientes con un verbo transitivo y un

sujeto presentes en los datos. Es de notar que la topicalidad del referente también motiva la posición preverbal del sintagma nominal en otros tipos de construcciones (Ocampo 2002, 2004, 2005a, 2005b, 2009, 2013). El análisis de los datos indica que no hay correlación entre foco y orden de las palabras (Ocampo 2003a, 2003b, 2013). La posición tradicional considera que la prominencia prosódica coincide con el foco y que tópico y foco están en oposición. El tópico coincide con la información presupuesta (información conocida) y el foco con la información no presupuesta (no conocida) presente en la oración. La información no presupuesta (es decir el foco) aparece al final de la oración y en castellano está marcada por medio de la prominencia prosódica que cae en la sílaba acentuada de la última palabra. Toda prominencia prosódica que no caiga allí es contrastiva o enfática (Zubizarreta 1999). El problema es que esta hipótesis se basa en datos introspectivos, creados por el analista, que solamente revelan posiciones prototípicas cuando son correctos (Ocampo 2010). Si se analizan datos reales, es decir tomados de conversaciones informales, el panorama cambia. El análisis de estos datos permite hipotetizar que el foco y el tópico no están en oposición sino que pertenecen a ámbitos distintos. El tópico es la manifestación sintáctica del mecanismo cognitivo descubierto por Morton Ann Gernsbacher (Gernsbacher 1997, Gernsbacher y Hargreaves 1992) *the structure building framework* (Ocampo 2003b). Por otra parte, en los datos conversacionales, la prominencia prosódica no contrastiva no cae necesariamente en la porción no presupuesta de la oración ni en el último constituyente. En realidad, los datos revelan que cualquier parte de la oración puede recibir prominencia prosódica no contrastiva (Ocampo 2003a). La prominencia prosódica principal, es decir el foco, es una instrucción del hablante a la audiencia para que centre la atención en el constituyente sobre el cual cae y haga las inferencias adecuadas. Esta marcación, de índole pragmática, del centro de atención no tiene influencia directa en el orden de las palabras de la oración. Como puede observarse en los ejemplos (1), (2), (5), (6), (7), no hay ninguna correlación entre el centro de atención –el foco– y el orden de palabras.

Los factores que afectan el orden de las palabras en las oraciones simples e independientes pueden agruparse en dos categorías: factores cognitivo-sintácticos y factores pragmáticos. (Ocampo 2013). Los factores cognitivo-sintácticos incluyen varios tipos de componentes de origen cognitivo que se han codificado en la sintaxis. Estos factores comprenden: tipo de construcción, número de constituyentes, topicalidad de los referentes de los

sentencias nominales. Los factores pragmáticos comprenden: constituyente contrastivo, constituyente realzado, hipótesis, contrario a lo esperado, etc. Hay una correlación entre ambos tipos de factores (Ocampo 2013), que por razones de espacio y no ser totalmente relevante a la presente exposición no se expone aquí.

Es significativo que en las cláusulas relativas presentes en los datos, el tópico sea el único factor que motiva el orden de las palabras. Es decir, que solamente están presentes factores de tipo cognitivo-sintáctico<sup>5</sup>. Si se contrastan estos resultados con los factores que motivan el orden de las palabras en oraciones independientes con un verbo transitivo y un sujeto, se observan diferencias. Los factores y las frecuencias se presentan en el Cuadro 3.

Orden de palabras	Motivación	Cantidad	Frecuencia
(a) Sujeto Verbo	Sujeto tópico	63/83	75.90
(b) Sujeto Verbo	Sujeto tópico contrastivo	5/83	6.03
(c) Verbo Sujeto	Sujeto baja topicalidad	13/83	15.66
(d) Verbo Sujeto	Verbo contrastivo	2/83	2.41

**Cuadro 3.** Orden de palabras, motivación y frecuencias en oraciones independientes con un verbo transitivo y un sujeto

El Cuadro 3 muestra que en las oraciones independientes con un verbo transitivo y un sujeto, los factores que motivan el orden de las palabras, además de topicalidad son: sujeto contrastivo o verbo contrastivo. Por consiguiente, además de los factores cognitivo-sintácticos (topicalidad) aparecen factores pragmáticos (constituyente contrastivo). Por motivos de espacio y teniendo en cuenta que el tema del trabajo no lo constituyen las oraciones simples sino las relativas y que las oraciones simples sólo se presentan como comparación, los factores presentados en el Cuadro 3 se ejemplifican sumariamente a continuación:

<sup>5</sup> Los otros dos factores cognitivo-sintáctico que están inherentemente presentes en las relativas son: tipo de construcción (cláusulas relativas con un sujeto y un verbo transitivo) y número de constituyentes (dos).

- (8) 9a15  
M: Entonces se va a la casa de esta señora y le dice que venía a bus buscar el televisor que había hablado para que se lo arreglara () /y la seño'ra se lo entre'ga/.
- (9) 23a12  
M: Y otra p cosa, la fuente de trabajo que tenían los - la gente: m - gente m (0.4) carenciada, muy humilde, que era (0.2) que son los puestos que están en la calle sie:te, en la calle o:cho, en diagonal oche:nta, en la calle doce, /el o'tro gobie''rno no los deja'ba./ /E'ste los de'ja./
- (10) 9b8  
M: pero era buenísimo () no no me acuerdo de ningún perro [...] que haya habido en casa bueno como ese  
L: no: ninguno  
M: ninguno como ese  
L: y: () y qué sentimiento  
[  
A: ¿y cómo fue que se murió?  
M: lo pisó un auto en la calle  
[  
L: /bue'no/, **lo mató un au'to**/.
- (11) 25b12  
S: bueno, cambió un poco la enseñanza, porque antes e:: () la maestra estaba: () suger - sugería más, [...] ahora se deja más que el chico reaccione () que que surja más del chico que que impone a la maestra. [...] Pero ahora no, () ahora el chico - o el chico tiene que llevar la caja para: para hacer un autito () o: si están dando plantas tienen que llevar plantitas para que la ma - () lleva () /**apo'rta el chi'co**/ () y no del: - de la escuela para el chico.

El ejemplo (8) muestra el factor (a) del Cuadro 3, el sujeto es de alta topicalidad y aparece en posición preverbal; (9) ejemplifica dos construcciones con sujetos de alta topicalidad y contrastivos: la hablante contrasta el gobierno anterior al actual, que es el factor (b) del cuadro 3; (10) ejemplifica (c) del Cuadro 3, una construcción con un sujeto de baja topicalidad en posición postverbal: el tópico de la construcción es el referente del perro; finalmente (11) muestra un caso de verbo contrastivo

con sujeto postverbal, (d) en el Cuadro 3. Aquí, el tópico de la oración es el referente del sintagma nominal sujeto *el chico*. El sujeto aparece en posición postverbal porque el verbo es contrastivo, la hablante por un lado autocorrigió el verbo, lo cual es un caso de contraste, pero además contrasta las acciones que ahora debe realizar el alumno con las que hacía la escuela antes.

Como ya se expresara anteriormente, en las cláusulas subordinadas relativas hay siete construcciones fragmentadas de las cuales no fue posible establecer el grupo melódico. De las seis construcciones con sujeto preverbal, dos son fragmentadas. Lo interesante es que en las cuatro construcciones restantes, el grupo melódico abarca solamente la cláusula relativa, como puede verse en (1) y (5). Por otra parte, en la mayoría de las construcciones con sujeto postverbal, el grupo melódico abarca la cláusula principal y la subordinada (2), (6) o el sintagma nominal y la cláusula subordinada que lo modifica (7). Las frecuencias se presentan en el Cuadro 4<sup>6</sup>.

Grupo melódico	Cantidad	Frecuencia
X /que S V/	4/4	100
/X que V S/	24/31	77.42
X /que V S/	7/31	22.58

**Cuadro 4.** Grupo melódico y orden de palabras

El Cuadro 4 muestra que la totalidad de las relativas con sujeto preverbal presentan un grupo melódico que comprende solamente la relativa, mientras que en la mayoría de las construcciones con sujeto postverbal 24/31, 77.42%, el grupo melódico abarca por lo menos el sintagma nominal antecedente de la relativa. Solamente 7/31, 22.58% de las cláusulas relativas con sujeto postverbal presentan un grupo melódico que abarca solamente a la relativa. Este caso se ejemplifica en (12):

(12) 7a1

P: me preguntó de: una: lámpara que teníamos nosotros

F: mhm

<sup>6</sup> Como se expresara anteriormente, del total de 42 construcciones relativas hay 7 fragmentadas, en las que no se determinó el grupo melódico.

P: que si lo íbamos a poner o no? yo le dije que no que lo use  
 () para ponerlo en el living

[

E: qué lámpara?

P: /e"sa de mi:'mbre, /**que nos dio**^ **tu mamá**'/.

El tópico de la relativa es el referente de *la lámpara de mimbre*. El sujeto postverbal *tu mamá* es de baja topicalidad.

Una correlación con el orden de palabras que aparece en los datos es que la totalidad de las cláusulas relativas con sujeto preverbal son explicativas, como puede observarse en (1) y (5), mientras que la mayoría de las relativas con sujeto postverbal son especificativas, como lo ejemplifican (2) y (6). Nótese que (12) es uno de los pocos casos de una relativa explicativa con sujeto postverbal.

Podría argumentarse que en (5) la subordinada es una relativa especificativa porque es la continuación del pronombre relativo anterior *las que [...] papá había usado*. Si así fuera, esta interpretación sería correcta. Pero no es esto lo que la hablante hace. Ésta se autocorrigió y vuelve a expresar el referente de *zapatillas* mediante un sintagma nominal formado por un adjetivo sustantivizado *las últimas*. La hablante destaca la cualidad de último con respecto al referente: nótese que el SN recibe el más alto grado de prominencia prosódica seguido de una entonación de coma que lo separa del resto de la emisión. En este contexto la hablante hace que la información transmitida por la relativa se vuelva menos crucial; es una información que se añade al elemento central constituido por *las últimas [zapatillas]*. Nótese que durante la conversación los hablantes dan instrucciones a los oyentes para que hagan determinadas inferencias. La relativa está enmarcada por dos entonaciones de coma, que son instrucciones de la hablante a la audiencia para que la interpreten como añadiendo información al referente *zapatillas*.

Hay una correlación entre el tipo de relativa —especificativa, explicativa— y el grupo melódico: el grupo melódico de las especificativas abarca por lo menos el sintagma nominal antecedente, mientras que las relativas explicativas tienen su propio grupo melódico. Por lo tanto, volviendo al Cuadro 4 se puede concluir que, en los datos, las construcciones con sujeto preverbal son relativas explicativas y tienen un grupo melódico propio, mientras que las construcciones con sujeto postverbal son en su mayoría, 77.42%, relativas especificativas con un grupo melódico que abarca por lo menos el sintagma nominal antecedente del pronombre relativo.

Es de consignar que la correlación entre grupo melódico y especificativa versus explicativa ya había sido notada por lo menos desde Navarro Tomás (Brucart 1999:397 nota 2); Givón (1990:645) describe la misma correspondencia para el inglés.

### **3. ANÁLISIS DE LOS RESULTADOS**

Teniendo en cuenta los resultados expuestos en la sección anterior se pueden plantear varias preguntas. El Cuadro 1 registra el pequeño número de cláusulas relativas con un verbo transitivo y un sujeto presentes en los datos, en comparación con el mismo tipo de construcción en las oraciones independientes que aparecen en el corpus. Se puede indagar entonces acerca de la causa de esa diferencia. El Cuadro 3 consigna la presencia de factores pragmáticos causantes de una variación en el orden de palabras, además del factor cognitivo-sintáctico de topicalidad. ¿Por qué, entonces, en las relativas el único factor presente que afecta el orden de las palabras es la topicalidad del sujeto? Es decir ¿por qué solamente aparece un factor cognitivo-sintáctico y los factores pragmáticos no aparecen? Por último, cabe preguntarse si hay alguna motivación de la relación entre orden de palabras, el tipo de cláusula relativa —especificativa o explicativa— y los dos tipos de grupo melódico presentes.

Con respecto a la primera pregunta, la respuesta no es definitiva por falta de más datos empíricos. Varios analistas han mencionado que en la lengua oral las cláusulas subordinadas son menos frecuentes. Esto es sólo una observación general que no indica la motivación. Son necesarios, como se acaba de expresar, más datos para poder establecer una relación de causa-efecto.

El menor número de factores promotores de variación en el orden de las palabras en las cláusulas relativas ha sido considerado por varios analistas, entre ellos Givón (1979) y Bybee (2001) quienes postulan la siguiente hipótesis, que considero satisfactoria: “Subordinate clauses contain backgrounded information that is less likely to be subject to topicalization, contrast and presentative focus; such manipulations are more appropriate and more commonly occur in main clauses” (Bybee 2001:2). Como se expresó antes, los factores sintáctico-pragmáticos están de alguna manera codificados en la lengua. Es posible, entonces postular que el ser ya parte de la gramática es un factor que facilita la aparición de la topicalidad en las



relativas, mientras que el hecho de no ser parte de la gramática hace que la presencia de los factores pragmáticos sea menos frecuente y por eso no aparezcan en los datos.

Con respecto a la tercera pregunta, varios analistas han notado la mayor autonomía sintáctica de las cláusulas explicativas en relación con las especificativas (Lehmann 1998, Brucart 1999 quien también cita a Bello). Brucart (1999) expresa que las relativas explicativas tienen elementos en común con la aposición y que su comportamiento se asimila al de los complementos adjuntos. Esta mayor autonomía sintáctica se correlaciona con la presencia de un grupo melódico independiente. Nótese la observación de Givón (1990:680): "The presence of an intonation break makes for a less defined, more permissive syntactic environment." Otro factor a considerar en esta interpretación lo constituye la reflexión de Chafe (1987), quien provee una explicación cognitiva de la existencia del grupo melódico: un grupo melódico es la expresión oral de la limitada cantidad de información que la mente humana es capaz de mantener en la conciencia en un momento determinado.

Basándose en las consideraciones anteriores se puede hipotetizar que la manifestación de sujetos preverbales solamente en cláusulas explicativas se debe a que estas cláusulas tienen mayor autonomía sintáctica, lo que les permite expresar una idea separada de la expresada por la principal, en este caso un tópico independiente del tópico de la cláusula principal. Esto a su vez tiene correlato con la presencia de un grupo melódico independiente. Lo anterior lleva a la consideración que las subordinadas explicativas no son relativas prototípicas. Se puede postular entonces que las relativas prototípicas son las especificativas, donde no se comunica una idea independiente de la expresada en la principal. Esto motiva un orden de palabras con sujeto postverbal, ya que el tópico de la subordinada coincide con el de la cláusula principal, y motiva también la ausencia de un grupo melódico propio. Por consiguiente, el centro del prototipo lo ocupan las relativas especificativas, con sujeto postverbal sin grupo melódico independiente. Las relativas especificativas con sujeto postverbal y grupo melódico independiente ocupan una posición menos central, mientras que las relativas explicativas con sujeto preverbal y grupo melódico independiente ocupan una posición marginal.

#### **4. RESUMEN Y CONCLUSIÓN**

En este trabajo he tratado de presentar una motivación de la variación del orden de palabras presente en las cláusulas relativas con un verbo transitivo y un sujeto. La motivación primera de la variación, la topicalidad del referente, también está presente en construcciones independientes, pero a esto hay que sumarle la presencia de otros factores en las relativas que se correlacionan con esta variación.

Es de notar que el descubrimiento de esta íntima interrelación de factores ha sido posible por la utilización de un corpus de datos de conversaciones informales, en oposición a construcciones creadas por el analista, las cuales no permiten descubrir la vasta gama de elementos presentes en este tipo de variación (Ocampo 2010).



## REFERENCES

- Bentivoglio, Paola (1985): "Función y significado de la postposición del sujeto nominal en el español hablado". Caracas: Instituto de Filología Andrés Bello, Universidad Central de Venezuela.
- y Elizabeth Weber (1986): "A functional approach to subject word order in spoken Spanish", en Osvaldo Jaeggli y Carmen Silva-Corvalán (eds.): *Studies in Romance Linguistics*, Amsterdam: Floris, pp. 23-40.
- Brucart, José María (1999): "La estructura del sintagma nominal: las oraciones de relativo", en Ignacio Bosque y Violeta Demonte (eds.): *Gramática descriptiva de la lengua española*. Madrid: Espasa-Calpe, pp. 394-522.
- Bybee, Joan (2001): "Main clauses are innovative, subordinate clauses are conservative. Consequences for the nature of constructions", en Joan Bybee y Michael Noonan (eds.): *Complex sentences in grammar and discourse. Essays in honor of Sandra A. Thompson*. Amsterdam/Philadelphia: John Benjamins, pp. 1-18.
- Gernsbacher, Morton Ann (1997): "Two decades of structure building", *Discourse Processes*, 23:3, pp. 265-304.  
<http://dx.doi.org/10.1080/01638539709544994>
- y David Hargreaves (1992): "The privilege of primacy, experimental data and cognitive explanations", en Dorys Payne (ed.): *Pragmatics of Word Order Flexibility*. Amsterdam/Philadelphia: John Benjamins, pp. 83-116.
- Chafe, Wallace (1987): "Cognitive constraints on information flow", en Rusell Tomlin (ed.): *Coherence and grounding in discourse*. Amsterdam/Philadelphia: John Benjamins, pp. 21-51.

- (1994): *Discourse, consciousness, and time. The flow and displacement of conscious experience in speaking and writing*. Chicago: The University of Chicago Press.
- Givón, Talmy. (1979): *On understanding grammar*. New York: Academic Press.
- (1990): *Syntax. A functional-typological introduction*. Volume II. Amsterdam/ Philadelphia: John Benjamins.
- Cruttenden, Alan (1986): *Intonation*. Cambridge: Cambridge University Press.
- Kahane, Henry y Renée Kahane (1950): “The position of the actor expression in colloquial Mexican Spanish”, *Language* 56.2.
- Lehmann, Christian (1988): “Towards a typology of clause linkage”, en John Haiman y Sandra A. Thompson (eds.): *Clause combining in grammar and discourse*. Amsterdam/ Philadelphia: John Benjamins, pp. 181-225.
- Levinson, Stephen (1983): *Pragmatics*. Cambridge: Cambridge University Press.
- Ocampo, Francisco (2002): “The word order of constructions with *ser* and *estar*, a subject NP and an adjective in spoken Spanish”, en James Lee, Kimberly Geeslin y Clancy Clemens (eds.): *Structure, Meaning, and Acquisition in Spanish. Papers from the 4<sup>th</sup> Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Press, pp.212-229.
- (2003a): “On the notion of focus in spoken Spanish: An empirical approach”, en Paula Kempchinsky y Carlos-Eduardo Piñeiros (eds.): *Theory Practice, and Acquisition. Papers from the 6<sup>th</sup> Hispanic Linguistics Symposium and the 5<sup>th</sup> Conference on the Acquisition of Spanish and Portuguese*. Sommerville, MA: Cascadilla Press, pp. 207-226.
- (2003b): “The expression of topic in spoken Spanish. An empirical study”, en Rafael Núñez-Cedeño, Luis López y Richard Cameron (eds): *A Romance perspective on language knowledge and use*. Amsterdam/ Philadelphia: John Benjamins, pp. 195-208.
- (2004): “Word order variation in spoken Spanish in constructions with a verb, a direct object, and an adverb. The interaction of syntactic, cognitive, pragmatic, and prosodic features”, en Robert Kirsner, Ellen Contini-Morava y Betsy Rodriguez-Bachiller (eds.): *Cognitive and Communicative Approaches to Linguistic Analysis*. Amsterdam/Philadelphia: John Benjamins, pp. 341-360.

- (2005a): “The Word Order of Constructions with an Intransitive Verb, a Subject, and an Adverb in Spoken Spanish”, en David Eddington (ed.): *Selected Proceedings of the 7<sup>th</sup> Hispanic Linguistics Symposium*. Sommerville, MA: Cascadilla Press, pp. 142-157.
  - (2005b): “Variación en el orden de palabras: construcciones con un verbo intransitivo, un sujeto y un adverbio”, en Angelita Martínez (ed.): *Signo y Señal. Variación 13*. Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Argentina, pp. 255-270
  - (2009): “El orden de palabras en el español hablado. La construcción Sujeto Verbo Objeto Directo”, en Montserrat Veyrat Rigat y Enrique Serra Alegre (eds.): *La lingüística como reto epistemológico y como acción Social. Estudios dedicados al Profesor Ángel López García con ocasión de su sexagésimo aniversario*, Volumen 1. Madrid: Arco Libros, pp. 501-511
  - (2010): “The place of conversational data in Spanish syntax: Topic, focus, and word order”, *Studies in Hispanic and Lusophone Linguistics*, volume 3, issue 2, pp. 533-543.
  - (2013): “No, pero siguen las supersticiones. Sí, las supersticiones siguen. El orden de palabras en rioplatense: factores cognitivos, sintácticos y pragmáticos”, en Laura Colantoni y Celeste Rodríguez (eds.): *Perspectivas teóricas y experimentales sobre el español de la Argentina*. Madrid: Iberoamericana Verbuert, pp. 299-319.
- Silva Corvalán, Carmen (1977): *A discourse study of some aspects of word order in the Spanish spoken by Mexican-American in West Los Angeles*. Tesis de Maestría, University of California Los Angeles.
- (1981): “The diffusion of object-verb agreement in Spanish”. *Papers in Romance*, 8.2, pp. 163-176.
  - (1983): “On the interaction of word order and intonation: Some OV constructions in Spanish”, en Flora Klein-Andreu (ed.): *Discourse perspectives on syntax*. New York: Academic Press, pp. 117-140.
  - (1984): “Semantic and pragmatic factors in syntactic change”, en Jacek Fisiak (ed.): *Historical Syntax*. Berlín/New York: Mouton, pp. 555-573.

Sosa, Juan Manuel (1999): *La entonación del español. Su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra.

Zubizarreta, María Luisa (1999): “Las funciones informativas: tema y foco”, en Ignacio Bosque y Violeta Demonte (eds.): *Gramática descriptiva de la lengua española*, vol 3. Madrid: Espasa-Calpe, pp. 4215-4243.

## Apéndice:

### CONVENCIONES UTILIZADAS EN LA TRANSCRIPCIÓN

- Los números y la letra que aparecen en cada ejemplo indican su ubicación en el corpus.
- Un corchete '[' entre dos turnos de habla indica que se emiten de manera simultánea.
- El signo '=' entre una emisión y la siguiente indica que hay una pequeña superposición entre el final de un turno y el comienzo del siguiente.
- Las emisiones aparecen entre paréntesis cuando la precisión de lo transcripto no es segura. Tres puntos entre paréntesis '(...)' indican que fue imposible transcribir lo que dice el hablante.
- Tres puntos entre corchetes '['...]' indican que, por razones de espacio, se omitió parte de la transcripción en el ejemplo.
- Toda interrupción en el habla se indica con un guión '-'.
- Una pausa se indica por medio de paréntesis '()'. Si hay números entre paréntesis, estos indican la duración de la pausa en segundos y décimas de segundo '(1.4)'.
- Los trozos subrayados indican que son salientes (acento primario, tono alto, alta intensidad, etc.).
- El alargamiento vocálico o consonántico se indica por medio de dos puntos ':' después de una vocal o consonante.
- Un punto indica entonación de punto; una coma, entonación de coma.
- Las barras inclinadas '/' indican grupo melódico.



- El acento primario se indica con ' , el acento secundario con " y el acento terciario con ^.
- Toda información adicional aparece entre corchetes '[risas]'.
- Los nombres que aparecen en las conversaciones fueron cambiados para preservar el anonimato. Los nombres de personalidades televisivas o políticas no se cambiaron. El nombre del investigador, cuando aparece, no se cambió.
- La porción relevante del ejemplo aparece **en negrilla**.

## Cláusulas flexionadas predicativas con verbos de percepción

María José Rodríguez-Espiñeira  
*Universidad de Santiago de Compostela*

**Abstract:** This paper proposes a discursive presentational analysis of the sentence type *Vio a los gemelos Vicario que corrían hacia la iglesia* (“She saw the Vicario twins running towards the church”) by adapting a hypothesis from Lambrecht (2000) for French. It is argued that this construction is pragmatically motivated and that it shares discursive properties with the so-called proleptic constituents (*La descubrí que me fue infiel* “I found out that she cheated on me”). It is claimed that, in the most typical cases, the event expressed in the finite subordinate clause allows to reveal or to confirm the identity of the object referent. Corpus data from CREA clearly show that there exist interferences between finite clauses behaving as secondary predications —denoting direct sensorial perception or no epistemic perception— and object complement clauses —denoting typically epistemic perception, on sensorial grounds or purely cognitive—. Blending takes place when the theme argument in pronominal form is signaling a given discourse entity that functions as a reference point for the perception event.

**Keywords:** Spanish perception verbs, finite clauses, secondary predication.



## 1. INTRODUCCIÓN

Este trabajo examina las propiedades formales, semánticas y discursivas de la construcción ejemplificada en (1a), documentada en varias lenguas románicas (cf. Auwera 1985, 1993, Rafel 1999 y referencias allí citadas). La construcción ha recibido diversas etiquetas –pseudo-relativa, reseña de percepción, relativa presentativa, entre otras– debido a los debates existentes sobre el carácter del nexa *que*, relativo o conjunción, y sobre el carácter de la cláusula flexionada, relativa (1a.i) o completiva sustantiva (1a.ii). En la mayor parte de los trabajos sobre el tema, las propiedades de la construcción se ilustran con ejemplos creados *ad hoc* por los investigadores y, por regla general, carentes de contexto discursivo, lo que impide examinar algunas hipótesis que tratan de explicar las diferencias semánticas y discursivas entre esa variante —en cualquiera de los análisis de i y ii— y la de (1b).

- (1) a. (Oyó la voz de Estrella y cuando Martín se enderezó) **la vio que se ponía de pie**
- i. La<sub>OBJETO DIRECTO</sub> vio [que<sub>SUJETO</sub> se ponía de pie]<sub>CLÁUSULA RELATIVA PREDICATIVA</sub>
- ii. La<sub>EXTRAPOSICIÓN</sub> vio [que se ponía de pie]<sub>OBJETO DIRECTO</sub>
- b. vio [que {ella/Ø}<sub>SUJETO</sub> se ponía de pie]<sub>OBJETO DIRECTO</sub>

El objetivo de este trabajo es aplicar a los datos de corpus recopilados<sup>1</sup> un tipo de análisis de su estructura informativa en la línea de Lambrecht (2000) para el francés. Según este autor, las denominadas relativas de percepción son construcciones presentativas, pese a que el antecedente es una entidad identificable (un nombre propio, una frase nominal de referente específico) o ya identificada discursivamente (un clítico). Cuando un verbo de percepción se combina con una relativa predicativa, una entidad es presentada o recontextualizada en un tramo concreto del discurso por mediación del acto de percepción —de ahí su carácter presentativo— y en la

---

<sup>1</sup> La muestra de trabajo está constituida por 130 ejemplos con complemento directo en forma de clítico extraídos de los corpus CREA y CORDE y por 66 ejemplos con complemento directo en forma de frase plena, provenientes del CREA. Los datos son resultado de las siguientes consultas: a) clítico + verbo + *que*, con diferentes personas y tiempos verbales; b) verbo + *a* + *que*, con diferentes personas y tiempos. Para acotar las búsquedas se ha utilizado el operador de distancia DIST/, tal y como se explica en el §3.4.3 de la Ayuda de los bancos de datos. También se han efectuado consultas puntuales en GOOGLE para confirmar o descartar algunas opciones. Este trabajo forma parte del proyecto FFI2012-38550 (MINECO y Fondos FEDER).

cláusula subordinada se hace un comentario sobre esa entidad. Lo destacado en (1a) es que un individuo concreto entra en el campo visual del sujeto perceptor al tiempo que realiza un movimiento.

No figura entre las metas de este estudio discutir el carácter del nexa, entre otros motivos porque tipológicamente suele admitirse que la relativización también opera cuando el nexa es un simple relator, sin función sintáctica en la cláusula que encabeza –lo que en la terminología inglesa se denomina *complementizer* (De Vries 2002)–. Para identificar un mecanismo relativizador se requiere que exista un elemento semánticamente compartido por la cláusula dominante y por la cláusula subordinada, un antecedente, al que llamaremos pivote, siguiendo a De Vries (2002: 14)<sup>2</sup>. En palabras de Silva-Corvalán (1999: 448), “el relativizador *que* enlaza dos cláusulas e indica que hay un elemento que guarda relación sintáctica con los verbos de las dos”. Y este requisito se cumple en el ejemplo de (1a).

La interpretación de la cláusula subordinada como relativa suele asociarse con otros postulados: a) El predicado regente de percepción selecciona tres argumentos, un perceptor, una entidad percibida –objeto directo codificado como frase nominal o expresión equivalente– y una cláusula relativa en función de complemento predicativo del objeto directo; b) El predicado subordinado satura su valencia en el seno de la cláusula flexionada –la forma *que* es a la vez nexa subordinante y sujeto de la cláusula subordinada– y el relativo tiene como antecedente el complemento directo del verbo regente. El esquema sintáctico que corresponde a esta interpretación se refleja en el análisis de (1a.i). Las relativas predicativas constituyen una forma de codificar un evento que actúa como predicación secundaria<sup>3</sup> y el verbo de percepción expresa percepción sensorial o percepción no epistémica<sup>4</sup>.

<sup>2</sup> La noción de antecedente se basa en la correferencialidad. La noción de pivote es ligeramente diferente, porque remite a un elemento compartido por la cláusula matriz y la relativa, que puede realizarse fonéticamente en la primera o en la segunda. Vid. también Pompei (2009: 428).

<sup>3</sup> Se entiende por secundaria una predicación –por regla general no finita o reducida– que está orientada hacia un participante central de un predicado primario (de percepción en nuestros datos), con el que guarda concomitancia temporal. En este trabajo asumimos que la función sintáctica de la predicación secundaria es la de ‘complemento predicativo’, en el caso analizado referido al objeto directo. Para cuestiones de terminología y análisis alternativos, vid. NGLÉ (2009: §26.9).

<sup>4</sup> Carrasco Gutiérrez (2011a) propone distinguir entre percepción no epistémica, percepción directa de individuos y de eventos, y percepción epistémica, percepción de contenidos proposicionales. Con la percepción no epistémica los hablantes cuentan lo que perciben y los referentes del complemento son personas, animales o cosas (*Vieron a Santiago/al gato/el barco*); también eventos en los que participan las entidades (*Vieron atracar el barco; a Santiago lo vieron subiendo al barco*). Mediante la percepción epistémica, los hablantes informan sobre conocimientos adquiridos. Dentro de este segundo tipo, Carrasco Gutiérrez distingue entre

En cambio, la consideración de la forma *que* como conjunción subordinante suele conllevar el análisis del predicado de percepción como biargumental: un perceptor y una situación o evento percibido, que funciona como objeto directo –la cláusula subordinada introducida por la conjunción *que*–. Las cláusulas sustantivas con predicados de percepción codifican preferentemente o bien la constatación intelectual de una percepción sensorial, o bien una percepción cognitiva, una deducción del perceptor basada en el razonamiento (cf. Hanegreefs 2008, Fernández Jaén 2012 y referencias allí citadas)<sup>5</sup>. Según la interpretación de (1a.ii) el sujeto de la cláusula subordinada sufre un proceso de extraposición (o *prolepsis*) que lo desplaza de su posición en el interior de la completiva (1b). Este fenómeno de transferencia suele analizarse en términos sintácticos: la variante de (1a.ii) se deriva de la de (1b), pero raramente se especifican los factores que la motivan o justifican (§4).

Como veremos en este trabajo, en algunos contextos las diferencias semánticas entre (1a) y (1b) son evidentes: percepción sensorial (‘percibir con el sentido de la vista’ en (1a)), frente a percepción epistémica (‘percibir con la inteligencia’ en (1b)). Pero en otros contextos, (1a) resulta una variable lingüística no conflictiva desde el punto de vista del significado referencial, pues lo comparte con (1b). En este trabajo se defenderá que, también en estos casos, existen diferencias de significado en el nivel pragmático o discursivo (§4 y 5).

Por otra parte, las cláusulas predicativas con verbos de percepción ofrecen cierto parecido superficial con las relativas explicativas, aspecto que han señalado muchos autores. Se distinguen de ellas por su diferente estructura prosódica<sup>6</sup> y también por su peculiar valor semántico e

---

percepción epistémica primaria (conocimiento conectado directamente con una experiencia visual) y percepción epistémica secundaria (conocimiento asociado a un proceso inferencial). Al decir *Miró por la ventana y vio que llovía* informamos sobre un conocimiento adquirido mediante constatación visual del evento de *llover*, pero si decimos *Vio que iba a llover* proporcionamos información obtenida a partir de datos secundarios, que deben ser interpretados por el perceptor (el cielo está cubierto, sopla un viento húmedo, suena un trueno, etc.).

<sup>5</sup> Hanegreefs (2008: 119) señala que la percepción física indirecta (codificada con cláusulas flexionadas) es la zona de transición gradual entre la percepción física directa de eventos (infinitivos y gerundios, preferentemente) y la percepción cognitiva (también codificada con cláusulas flexionadas). En una línea similar, Carrasco Gutiérrez (2011a) sostiene que las cláusulas flexionadas codifican tanto percepción epistémica primaria como secundaria.

<sup>6</sup> La relativa explicativa o no restrictiva, también llamada apositiva por tener forma de inciso, constituye un grupo fónico autónomo y la frontera que la separa del antecedente se marca con un tonema suspensivo o descendente (semicadencia). Por el contrario, el tonema que separa una relativa presentativa de su antecedente es ascendente (anticadencia o semianticadencia) (Brucart 1999: 430).

informativo: la relativa explicativa designa una propiedad del antecedente en forma de inciso, una modificación de naturaleza parentética de una entidad que es identificada de manera independiente en el discurso (2). El estatuto informativo del contenido proposicional de la relativa se presenta como carente de importancia y puede ser conocido o desconocido para el interlocutor (Lambrecht 2000). De sus características prosódicas, semánticas e informativas se deriva su función sintáctica: la relativa es un modificador explicativo que forma con su antecedente una frase nominal expandida –con la estructura<sub>FN</sub> [FN, *Claus rel.*]–, cuya distribución es la misma que la de una frase nominal ordinaria o no expandida. En cambio, la relativa predicativa de (3) funciona como complemento predicativo del complemento directo –*los gemelos Vicario y mi madre*, respectivamente– y la construcción muestra propiedades discursivas específicas, entre ellas la de hacer cognitivamente accesible una entidad merced al evento en el que está inmerso y mediante el que es percibida (vid. *infra* §§2 y 3 para un análisis más detallado).

- (2) a. En la comisaría de Nantes **conoció** a Constanza Martínez, que luego sería su mujer. [CREA. La Vanguardia, 27/03/1994]
- b. Abrí los ojos y **vi** a Omar, que me sacudía por los hombros [CREA. 1991. Quintero, Ednodio. *La danza del jaguar*. Venezuela]
- (3) a. Plácida Linero se asomó a la ventana de la plaza y **vio** a los gemelos Vicario que corrían hacia la iglesia. [BDS. Crónica, 117, 18]
- b. Por fin, **oigo** a mi madre que se despide y camina hacia mi cuarto [CREA. 1994. Urroz, Eloy. *Las plegarias del cuerpo*. México]

En (2a) el verbo *conocer* selecciona como complemento una entidad individual, expresada mediante el nombre propio *Constanza Martínez*, al que se añade un dato novedoso para el lector pero que el redactor presenta como secundario y poco relevante para la narración principal, centrada en mostrar el calvario de los deportados a campos de concentración. En (2b) aparece un verbo de percepción visual y, dadas las características de la cláusula relativa explicativa, el movimiento de *sacudir los hombros* podría resultar relevante para la percepción; sin embargo, en el texto se ha separado el antecedente y la relativa con una coma, lo que dota a la cláusula de carácter parentético. La ausencia de pausa en (3) invita al lector a interpretar que

la predicación secundaria es el elemento desencadenante de la percepción visual o auditiva de la entidad percibida. La puntuación se revela, pues, como un factor decisivo (vid. nota 7).

En los próximos apartados analizaremos con cierto grado de detalle las cláusulas predicativas con objeto de diferenciarlas de otras construcciones próximas. En el apartado 2 revisaremos las propiedades formales, semánticas y funcionales de la construcción y en el apartado 3 nos ocuparemos de los aspectos informativos. Dedicaremos el apartado 4 a analizar ejemplos en los que aparece un clítico –con antecedente textual y sin él–, ya que los datos sugieren que bajo una misma apariencia formal coexisten dos construcciones específicas que difieren en cuanto a valor gramatical y discursivo. Por último, en el apartado 5 cotejaremos las cláusulas predicativas con cláusulas completivas en función de objeto directo. El trabajo se cierra con un apartado de conclusiones.

## 2. ASPECTOS FORMALES, SEMÁNTICOS Y FUNCIONALES

Una cláusula predicativa puede diferenciarse, al menos formalmente, de una relativa explicativa o incidental, porque carece de las propiedades entonativas asociadas a la pausa, representada con una coma en la lengua escrita. Tampoco es analizable como relativa especificativa, porque no restringe la referencia de su potencial antecedente. En el par de (4) se comparan una relativa restrictiva (4a) y una relativa predicativa (4b), ambas bajo el dominio de un verbo de percepción:

- (4) a. ¿Ves ese coche rojo que va delante de la furgoneta?  
[CREA. 2001. Beccaria, Lola. *La luna en Jorge*. España]  
b. María cierra los ojos y ve a su hermana Alix que le dice  
antes de fugarse: Espera unos días y vengo por ti. [CREA.  
2002. Mendoza, Mario. *Satanás*. Colombia]

En (4a) la localización relativa del coche proporciona, junto con el color, un dato esencial para distinguir el referente del objeto percibido; ambos rasgos facilitan la identificación del referente al margen del tipo específico de predicado primario o dominante, de modo que podríamos trasladar la frase a otra secuencia, como *¿Te gusta ese coche rojo que va delante de la furgoneta?* La frase nominal de (4a) contiene el demostrativo *ese*: los deícticos permiten una identificación del referente externa al enunciado, vinculada a la propia situación comunicativa, lo cual contradice las características de



las relativas predicativas (Kleiber 1988: 110). En cambio, en (4b), la entidad percibida posee un referente inequívoco (*su hermana Alix*) y la entidad que experimenta la percepción (*María*) ha cerrado los ojos, lo que impide la percepción sensorial, por lo que el acto de habla que contiene la cláusula predicativa (*le dice antes de fugarse...*) es precisamente lo que le permite a María recrear una imagen previa y revivir el momento en que su hermana le promete regresar para llevarla consigo.

Las cláusulas predicativas están restringidas a un subconjunto muy preciso de predicados regentes, en concreto verbos de percepción y de descubrimiento. Se documentan principalmente con los verbos *ver* y *oír*, son más raras con *escuchar* (5a) y con *mirar* —con este verbo se usan en acotaciones teatrales (5b)<sup>7</sup> o en fragmentos narrativos muy concretos— y se registran esporádicamente con verbos como *sentir* (6a), o *descubrir* (6b)<sup>8</sup>:

- (5) a. [...] y cerrando los ojos **escuché** a Carlitos Gardel que cantaba: “nada debo agradecerte, mano a mano hemos quedado...” [CREA. Los Tiempos, 20/11/1996. Bolivia]
- b. (SARAH se incorpora, **mira** a MOSS que oculta los prismáticos) [CREA. 1980. Vallejo, Alfonso, *Eclipse*. España. Teatro]
- (6) a. [el negro] Porque no se abalanzó contra él sino intentó escapar a través de él. Corrió y fue tan inesperado que Lituma no alcanzó a atajarlo y **lo sintió** que se estrellaba contra él [CREA. 1977. Vargas Llosa, Mario. *La tía Julia y el escribidor*. Perú]
- b. Ésta giró asustada y **descubrió** a Pedro que la miraba emocionado hasta las lágrimas. [CREA. 1989. Esquivel, Laura. *Como agua para chocolate*. México]

<sup>7</sup> Con este verbo, los textos de las obras teatrales ofrecen puntuaciones distintas. A veces aparecen cláusulas flexionadas no separadas por coma del pivote, como en (i) y otras toman la forma de un inciso, lo que impone una lectura como relativa explicativa, como en (ii):

(i) Excitado en medio del escenario mira a su padre que lo rehúye, a Patricia que baja la cabeza, y a Jenaro que se levanta y le sostiene la mirada. [CREA. 1990. Luca de Tena, María Luisa. *Un millón por una rosa*. España. Teatro]

(ii) Esteban mira a Shila, que tiene un estremecimiento, y comienza a canturrear una melopea. [CREA. 1992. Sanchis Sinisterra, José. *Nafragios de Álvar Núñez o La herida del otro*. España. Teatro]

<sup>8</sup> *Descubrir* se combina con predicaciones secundarias en forma de gerundio, no de infinitivo (*lo descubrió mirándola* / *\*mirar*). En todo caso, lo más probable es que los ejemplos con otros verbos, como *encontrar*, *descubrir*, *imaginar* o *sorprender*, mencionados en la bibliografía, se creen por analogía con los verbos de percepción (Lambrecht 2000: 61).

Uno de los principales argumentos para apoyar el análisis como complementos predicativos de estas cláusulas flexionadas es que equivalen funcionalmente a infinitivos, gerundios y participios que actúan como predicaciones secundarias —especialmente participios de presente, en las lenguas donde son funcionales—. Se crea así una clase paradigmática de elementos que conmutan en la función de predicativo del objeto directo (*las veo sonrientes/sonreír/sonriendo; los oyó discutir/discutiendo*). Pese a las diferencias aspectuales y de finitud de las cláusulas que forman esta clase paradigmática, en muchos contextos son intercambiables. En los fragmentos de (7) las cláusulas predicativas suceden a una predicación secundaria de gerundio (7a) y a dos de infinitivo (7b). En cada caso las predicaciones secundarias se refieren a las mismas entidades: Alberto Olaya en (7a) y Husan en (7b):

- (7) a. El caso de **Alberto Olaya** es otro, sé que se hubiera fajado conmigo en la misma clase. Pero no fue eso lo que me detuvo. Siento en su presencia que me rebasa con facilidad. **Lo vi un día hablando inglés con unos marineros**. Otro día pasé por donde él vive, y **lo vi que estaba jugando al ajedrez**. [CORDE. 1966. Lezama Lima, José. *Paradiso*. Cuba]
- b. Pudo **ver** a Husan **llegar a la otra orilla, sentarse en una roca** buscando el sol para escurrir rápidamente el agua, después **lo vio que caminaba a saltitos por aquellas callejas**, hasta perderse de nuevo siguiendo la misma margen del río [CREA. 1977. Lezama Lima, José. *Oppiano Licario*, Cuba]

Desde el punto de vista diacrónico se ha llegado a justificar la aparición de estas relativas predicativas como sustitutas de la predicación secundaria con participio de presente (Tekavčić 1980: §807-808)<sup>9</sup>, que era la construcción especializada en latín para expresar percepción directa de una entidad, ya que el esquema formal de Acusativo con Infinitivo (AcI) podía servir tanto para manifestar percepción sensorial como cognitiva (Maraldi 1980).

<sup>9</sup> Según este autor, la oración latina *vidi puellam colligentem flores* sería adaptada al italiano como *vidi una ragazza che coglieva fiori*. La construcción con relativa sería la sustituta del participio de presente, al perder este buena parte de sus usos en algunas lenguas romances. El francés, sin embargo, conserva el participio de presente con verbos de percepción —aunque con ciertas restricciones según Keliber (1988)— y también la construcción flexionada: *je le vois pleurant à chaudes larmes / je le vois qui pleurait*. Sobre la equivalencia funcional entre participios y cláusulas relativas en latín puede consultarse Pompei (1999: §1.5.1).

En los próximos epígrafes documentaremos con ejemplos de corpus las principales propiedades de las cláusulas predicativas, tanto si reflejan rasgos considerados prototípicos, como si ofrecen detalles que las alejan de ellos.

### 2.1.

La construcción admite nombres propios antes del nexo o bien la codificación de la entidad percibida mediante un pronombre átono en caso acusativo (6a, 7, *supra*). De ello suele deducirse que el objeto de percepción es un individuo, cuya participación en un evento lo hace perceptible, y esta interpretación puede inclinar la balanza hacia el análisis como relativa, ya que en las construcciones con subordinada sustantiva lo percibido es globalmente una situación o acontecimiento (Brucart 1999: 431). Con los verbos de percepción que llevan sujeto agentivo, como *mirar*, la codificación como subordinada sustantiva del percepto es marcada<sup>10</sup>, de modo que en (8) no es posible la sustitución de la cláusula predicativa por otra completiva (cf. \**Oswaldo mira que Yoly se acerca a sacarle los paquetes*):

- (8) Oswaldo la mira a Yoly que se acerca a sacarle los paquetes  
[CREA. 1986. Fernández Tisconia, Nelly. *Made in Lanus*.  
Argentina. Teatro]

### 2.2.

La construcción predicativa exige que el verbo de la subordinada exprese coincidencia temporal con el tiempo del predicado regente, al implicar la percepción física de un individuo en el momento de producirse el evento descrito en la cláusula flexionada. De ahí el contraste señalado por Brucart (1999: 431, nota 35) entre el enunciado *He visto a Luis que iba a la fiesta en su nuevo coche* y el improbable \**He visto a Luis que había ido a la fiesta en su nuevo coche*, ya que se requiere simultaneidad temporal entre el acto de percepción y el desplazamiento de Luis. Esa falta de coherencia temporal aparece, sin embargo, en algunos ejemplos de corpus muy similares

<sup>10</sup> En usos documentados con completiva (i, ii), parece producirse una neutralización de las oposiciones *ver / mirar, oír / escuchar* en el dialecto o idiolecto del hablante:

(i) “Por ejemplo, si en un hospital de día, los pacientes **escuchan que** en el gabinete se está cantando una determinada canción, algunos de ellos le pedirán al musicoterapeuta” [CREA. 1981. Benezon, Ronaldo O. *Manual de musicoterapia*. Argentina].

(ii) Infinitamente desconsolados **miran que** tienen que volverse a poner las mismas ropas [CREA. 1981. Cabada, Juan de la. *Pasados por agua*. México]

a los que contienen cláusulas predicativas, como el de (9a), que lleva un predicado regente en presente y otro subordinado en pretérito perfecto. En el cotexto previo el hablante alterna el verbo *ver* con *darse cuenta de*, lo que confirma una lectura epistémica del predicado de percepción (conocimiento reflexivo):

- (9) a. **Me he dado cuenta** que, coño, en aquel tiempo la cosa era uno más ¿qué te digo? más muchacho, uno no pensaba las cosas y ahora, hoy en día, yo veo que sí ha habido un cambio bastante. O sea, ¿está mejor? Si está mucho mejor, vale. ¿**Le parece** que ahora los jóvenes tienen mayor orientación que antes? Sí, o sea, las personas ¿qué te digo?, **los amigos míos** que se han criado aquí, esto y el otro, que han que pasaron por ahí, vale, yo **los veo que han tenido un cambio** [CREA. Oral. Venezuela]
- b. Te casarás y tendrás hijos y los amarás y **los verás partir de tu lado para vivir su propio destino** [CREA. 1983. Aguilera, Nuri. *La caricia rota*. México. Novela]

En contextos predictivos, como el de (9b), Kleiber (1988: 91) ha defendido para el francés que la combinación de dos futuros también está excluida con cláusulas predicativas finitas, pese a que no lo está con infinitivos. Los datos examinados para este trabajo confirman inicialmente esta tesis: no se han documentado predicaciones secundarias con verbo flexionado en futuro; el ejemplo de (9b) corrobora su empleo con infinitivo.

### 2.3.

Además de coincidente temporalmente, la cláusula flexionada debe ser imperfectiva (Auwera 1985: 226, Kleiber 1988: §2 y referencias allí citadas). En la bibliografía se ha defendido que solo las formas de presente y de imperfecto son admitidas en la construcción que analizamos, ya que muestran un evento en curso sin tener en cuenta sus límites inicial o final, lo que permite relacionarlas anafóricamente con los tiempos presente o pasado de la percepción (el verbo regente). Por eso, una forma perfectiva como *llegó* en (10b) incumple las propiedades de una predicación secundaria:

- (10) a. Ese domingo, repito, **la vi llegar** a casa ensimismada, le dije algo y no me respondió. [CREA. 1994. Urroz, Eloy. *Las plegarias del cuerpo*. México]

- b. Yo creo que eso tiene que ver con la experiencia que ella tenía. Porque yo **la vi que llegó muy bien**, y fácil. Y se pensionó [CREA, Oral, Costa Rica]

En (10a) el evento de *llegar a casa* tiene lugar en el mismo intervalo en el que se produce la percepción y el infinitivo simple no expresa perfectividad; el efecto de percepción concluida no procede del infinitivo, que es dependiente, sino de la forma personal *vi*<sup>11</sup>. En cambio, en (10b) ni existe dependencia temporal entre los dos pretéritos ni el verbo *vi* indica la percepción sensorial del personaje femenino, una tía de la narradora. En (10b), en consonancia con el predicado epistémico de la cláusula previa (*creo*), el verbo *ver* expresa un proceso mental, la constatación por parte del narrador de que su tía ha alcanzado con facilidad todas las metas académicas propuestas (profesora en un colegio, en un liceo y finalmente jubilada).

#### 2.4.

La cláusula predicativa describe un evento directamente percibido, por eso en la bibliografía, especialmente en la que analiza la construcción francesa, se ha descartado la posibilidad de que dicha cláusula indique una propiedad o describa un estado (Auwera 1985: 225; Lambrecht 2000: 58)<sup>12</sup>. La justificación es semántica: las propiedades no suelen asociarse a la simple captación de una entidad, sino que la cláusula lleva aparejada una carga evaluativa o interpretativa (*la veo gorda, relajada, contenta*), que no es propia de las construcciones analizadas<sup>13</sup>. Los estados transitorios o

<sup>11</sup> En la bibliografía esta forma de infinitivo suele parafrasearse por un imperfecto, *la vi que llegaba a casa ensimismada*. Para un estudio sobre las diferencias entre infinitivo e imperfecto en este tipo de combinaciones, vid. Kleiber (1988).

<sup>12</sup> Así, se ha interpretado como de dudosa aceptabilidad la secuencia *??je l'ai vu qui était petit*.

<sup>13</sup> La construcción predicativa con participio o adjetivo está frecuentemente vinculada con un empleo evaluativo del verbo de percepción. En el siguiente fragmento, hemos destacado en negrita tanto el predicativo adjetivo como la calificación de "apreciación" del propio narrador:

<sup>(i)</sup> "En lo artístico, la movida de Bahía es muy interesante y tengo entendido, además, que esta apreciación va más allá de los límites de la ciudad». Pero Bahía ciudad le deja tiempo para reflexionar:

«En otros aspectos **la veo triste**". [CREA. La Nueva Provincia, 03/04/1997. Argentina]

Con adjetivos que remiten a propiedades físicas, también se constata esta interpretación:

<sup>(ii)</sup> [...] me levanto y para mi sorpresa **encuentro** que el cielo **me parece rojo**, **veo azul la sangre**, etc."

[1995. García Suárez. *Qualia: propiedades fenomenológicas*. España]

No obstante, participios y adjetivos pueden expresar percepción sensorial directa:

<sup>(iii)</sup> En una escena veladamente dramática **se lo ve recostado sobre un diván ingravido**, en traje de baño, recibiendo rayos luminosos en el cuerpo, como si se hallara recostado en la arena de una playa". [CREA. 2001. Soublette, Gascón. *Mensajes secretos del cine*. Chile]

episódicos pueden percibirse sensorialmente (*la vio parada / desnuda*) pero es menos probable que desencadenen la percepción del objeto:

- (11) a. Pero la señora, yo **la veo que es chévere** y los hijos de ella también, o sea, las personas allá son muy distintas a las del barrio por aquí [CREA, Oral, Venezuela]
- b. [...] y el sol arde pero ella no se acordó de traerse traje de baño, y él le dice que se bañe sin nada, y paran, la chica se desviste detrás de unas rocas y se **le ve de muy lejos correr desnuda al mar**. Y ya después **se los ve que están tirados en la playa**, bajo las palmeras, ella con una especie de sarong hecho con la camisa de él, y él con los pantalones puestos, nada más, y descalzo, y no se sabe de dónde viene [CREA. 1976. Puig, Manuel. *El beso de la mujer araña*. Argentina]

El ejemplo de (11a) contiene el verbo *ver* con significado perceptivo-evaluativo y, el contenido de la cláusula flexionada se parece más al de una completiva de objeto. En cambio, en el fragmento narrativo de (11b), que simula reproducir verbalmente las escenas de una película, aparece un predicado estativo-situacional (*están tirados en la playa*) en una cláusula flexionada que tiene una distribución de constituyentes paralela a la cláusula de infinitivo predicativa empleada en la secuencia previa para visualizar la actividad de la protagonista.

El español contemporáneo ofrece dos opciones constructivas para predicados secundarios con carácter dinámico (infinitivo o gerundio) y parece haber optado por una codificación diferenciada para los predicados estativos: o bien adjetivos o bien participios. Las cláusulas flexionadas predicativas se usan preferentemente como sustitutas de los primeros, pero también alternan con los segundos, como se desprende del ejemplo (11b).

## 2.5.

La cláusula predicativa no contiene normalmente ni una negación ni un modal (cf. Auwera 1985: 226 y referencias previas allí citadas), carencias que también se han atribuido a las predicaciones secundarias de infinitivo y gerundio. Así, es excepcional el uso de combinaciones como

*Lo vi no subir las escaleras*<sup>14</sup>. Según algunos autores, este ejemplo sería admisible pragmáticamente en un contexto en el que el hablante percibiese los intentos de una entidad por realizar la actividad de *subir las escaleras* y renunciase a llevarla a cabo por culpa de algún impedimento<sup>15</sup>. En cuanto a nuestros datos, las cláusulas predicativas son afirmativas, salvo en dos ejemplos. En (12a) se ofrece una opinión del hablante sobre el referente del clítico (*el personal director y supervisor*). A su vez, (12b) encaja en los llamados casos de *prolepsis* o extraposición del sujeto de la subordinada (cf. *infra*, §4), al aparecer una frase nominal en posición temática con clítico correferencial; esta construcción resulta una alternativa discursiva a otra con completiva (*Veo que el Minoxidil no me mejora mucho mi pelo*). Ambos ejemplos coinciden en que el predicado primario no expresa percepción sensorial sino un proceso intelectual o cognitivo (‘juzgar, comprobar’):

- (12) a. De modo que la formación de conciencia en **el personal director y supervisor** y yo diría su mismo entrenamiento, a veces no están entrenados ni o sea **yo los veo que no** están ni preparados mentalmente, ni ni capacitados tampoco para ejercer el trabajo que a veces les exigen los programas [CREA. Oral. Venezuela. Formalidad baja]
- b. El Minoxidil **lo veo que no** me mejora mucho mi pelo  
[Google: <http://foro.recuperarelpelo.com/viewtopic.php?t=18125>]

## 2.6.

Con predicaciones secundarias, lo esperable en contextos negativos es negar que un participante sea percibido mientras interviene en un evento: *no lo vi subir las escaleras, no lo vi jugar, no lo veo andar con chicas, nunca lo vi doblársele las rodillas*. Cuando una cláusula flexionada

<sup>14</sup> De Google hemos extraído el siguiente ejemplo:

Su velocidad, habilidad, llevada, y fuerza en el remate que le ayudaron a realizar varias jugadas peligrosas en el partido. Todo eso se me olvida cada vez que *lo veo no pasar el balón a sus compañeros*, es como que él existiera solo. [<http://www.emelexista.com/analisis-y-calificacion-emelec-vs-nacional/>]

<sup>15</sup> Carrasco Gutiérrez (2011a) defiende que los eventos negativos pueden ser objeto de percepción no epistémica (sensorial): *Pedro estaba muy enfadado con María. La evitó varios días. No le cogió el teléfono. Y Juan lo vio llegar a no saludarla por la calle*. Según Fernández Jaén (2012: 323), es pragmáticamente viable un ejemplo como *Una vez vi a Luis no correr por culpa de un tirón*.



aparece en un contexto negativo, los datos indican que se usa una forma de subjuntivo (13) y no de indicativo; la negación se comporta, pues, como un inductor modal, al igual que ocurre cuando el complemento es una cláusula sustantiva (cf. *veo que tenemos otra alternativa / no veo que tengamos otra alternativa*):

- (13) a. Se pegaba mucho conmigo, iba a todos mis estrenos. Me dio temor: anda demasiado conmigo, **no lo veo que ande con mujeres**, a lo mejor no me salió bien [CREA. 1997. Paranaguá, Paulo Antonio. *Arturo Ripstein*. México]
- b. [Hablando de *mi papá*]. Y una de ellas fue la que **nunca lo vi aflojar** (como decíamos en el campo). Enfrentó situaciones muy difíciles siempre con los pies bien puestos en la tierra, **nunca lo vi que se le doblaran las rodillas**, a pesar del peso que llevaba en el espinazo [CREA. 1986 VV.AA. *Vida y palabra campesina*. Chile]

Lambrecht (2000: 59) considera inaceptables los ejemplos con predicados negativos y cláusulas predicativas porque, dada la función discursiva presentativa que adjudica a la cláusula predicativa, consistente en establecer la presencia de una nueva entidad en el campo perceptivo, negar la percepción supondría atentar contra el fin discursivo de la construcción. Volveremos sobre este aspecto en el §.4.

## 2.7.

Según Kleiber (1988: §2.3 ), frente a los infinitivos, las relativas de percepción no pueden expresar eventos repetidos ni aparecer en contextos frecuentativos, porque el evento expresado en la relativa no puede presentarse como válido para todos los intervalos de las situaciones de percepción correspondientes (*lo vi muchas veces, a veces, varias veces...*). El empleo de la relativa predicativa en este tipo de contextos equivaldría a afirmar que la cláusula es verdadera en cada una de las ocurrencias efectivas de percepción, lo cual es contradictorio con su valor informativo. De hecho en los corpus son recurrentes los ejemplos que contienen expresiones temporales que aluden a la inmediatez de la percepción, a su carácter repentino, como las de (14):



- (14) a. **De pronto** levanta la mirada y ve a Angélica que viene caminando hacia él [CREA. 2002. Mendoza, Mario. *Satanás*. Colombia]
- b. **En ese momento**, veo a mi vecino que pasa por ahí cerca, en su auto. Lo insulto [CREA. 2002. Bucay, Jorge. *El camino de la autodependencia*. Argentina]
- c. **En ese instante** lo vi que venía hacia mí. En el primer momento creí que fuera el jardinero, pero al oírlo llamarme, al oírle decir: “María”, lo reconocí en la voz... [CORDE. a 1896. Silva, José Asunción. *Traducción de cuentos de Anatole France*. Colombia]

Pero también se documentan ejemplos donde el contenido de la subordinada flexionada revela comportamientos habituales (15a) o se informa de que la percepción se repite (15b):

- (15) a. Y todo por no hacerme caso, porque ya yo le había dicho que tuviera cuidado con los trabajadores de la fábrica, que son unos irresponsables que nada más andan viendo a quién le hacen el chiste. Se lo dije cuando la vi que **andaba** con las trenzas muy peinadas y queriendo llevar recados a la fábrica. [CREA. 1990. Mastretta, Ángeles. *Arráncame la vida*. México.]
- b. Se pasa el día acariciando su úlcera de estómago y a su pastelera, Mady. **De vez en cuando** oye a Cosme Vila que habla por Radio Moscú [CREA. 1986. Gironella, José María. *Los hombres lloran solos*. España]

## 2.8.

La identificación de las cláusulas predicativas se ve favorecida por la codificación de la entidad percibida mediante una expresión referencial de interpretación específica: frase nominal definida, nombre propio, clítico. Es más conflictivo su reconocimiento cuando aparece una expresión referencial indefinida (*por la puerta a medio abrir vi a un viejo que tocaba el violín*), ya que es posible o bien una lectura de la cláusula flexionada como modificador restrictivo (*lo vi*, el viejo tocaba el violín habitualmente) o bien una lectura predicativa (*en ese preciso momento, lo vi que tocaba el violín*). Esta segunda interpretación implica admitir que el contenido de la cláusula predicativa es tratado como una información nueva para

el receptor (Lambrecht 2000: 53), con lo que la cláusula tiene función presentativa, como en (16)<sup>16</sup>:

- (16) A través del proceso de una conciencia carnal, Fanny Hill nos enseña la misma lección ética. En sus comienzos **vemos una moza campesina que llega a Londres**. Amanece el capitalismo. Se concentran en las grandes ciudades masas de hombres y mujeres, sin otra riqueza que sus brazos. La moza, engañada, entra en un prostíbulo. [CREA. El País, 11/09/1977]

El evento de *llegar a Londres* no parece constituir un rasgo identificador del referente independiente de la percepción –lo sería si la relativa fuese restrictiva–, sino un rasgo que contribuye a presentarla en escena (cf. *infra*, § 3).

## 2.9.

La marcación independiente de la entidad percibida como objeto (preposición *a*, clítico de acusativo, salvo en variantes con leísmo) se ve corroborada por los casos de pasivización, generalmente presentados para el francés: *Ton frère a été vu qui volait son voisin* (Auwera 1985: 224). En los corpus manejados para este trabajo no hemos localizado estas construcciones en español, que sí se registran tras búsquedas en Google:

- (17) a. “El móvil era el robo y según se estableció en la escena solamente **una persona fue vista que** andaba armada en la zona y que disparó contra los pasajeros”, dijo. Pese a un operativo de búsqueda del responsable, la policía todavía no reporta capturas. [<http://www.laprensagrafica.com/un-homicidio-y-un-asalto-hubo-donde-tirotearon-bus>]

<sup>16</sup> En el siguiente ejemplo, parecen convivir en el mismo fragmento una relativa restrictiva y dos relativas predicativas coordinadas. La interpretación depende de que la puntuación del fragmento refleje una emisión oral sin pausas antes del segundo *que*:

Te cuento: un día quedo en encontrarme con mi novia para ir al cine, pero llovía mucho, el agua por las rodillas no te miento, y entonces no viene mi novia porque vivía en Martínez, y te juro, en Corrientes había agua por todos lados. Yo me había caído y estaba totalmente mojado y **veo a una rubia que viene caminando por enfrente que se para y espera**. Me cruzo, y me acerco y le digo no sé... que llovía, qué se yo... [CREA. 1991. Polimeni, Carlos. *Luca*. Argentina]

- b. Al tener conocimiento de los hechos, los agentes investigadores iniciaron una búsqueda de dicha unidad de motor, **la cual fue vista que** transitaba hacia el centro de Tuxtepec, dándole alcance metros adelante [<http://www.oaxacadiaadia.com/?p=15977>]
- c. *Bernabé y Pancho, mis amigos acólitos, formaban parte de las travesuras y comíamos los mejores frutos. En la sacristía nos quedábamos a dormir y “vigilar” la fruta, pero en la noche sucedió algo muy extraño. Una figura oscura, como de una mujer, fue vista que salía de la notaría y cruzó toda la sacristía.* [[http://elvalencianomich.webs.com/apps/blog/show/prev?from\\_id=2329327](http://elvalencianomich.webs.com/apps/blog/show/prev?from_id=2329327)]

### 2.10.

Las cláusulas predicativas están sujetas a la misma restricción que infinitivos, gerundios y participios: el constituyente pivote es sujeto de la predicación secundaria y objeto del predicado principal. Esto justifica que el relator carezca de marca funcional, es decir, impide que el nexos lleve preposiciones introductorias, una carencia que suele aducirse para rechazar el carácter de cláusula relativa de la construcción predicativa<sup>17</sup>.

### 2.11.

En resumen, las predicaciones secundarias en forma de cláusula flexionada reúnen los siguientes rasgos: a) autonomía sintáctica de la entidad percibida; b) predicado regente que expresa percepción sensorial; c) predicado subordinado en indicativo, coincidente temporalmente con el regente, imperfectivo, que denota un evento dinámico, carente de negación o de auxiliar modal, y que expresa un evento singular; c) la cláusula subordinada tiene valor presentativo. Se ajustan a esta descripción los ejemplos de 3, 4b, 5, 6, 7, 8, 14, 16 y 17. Sin embargo, otros ejemplos se alejan de ese prototipo, porque el predicado subordinado indica anterioridad con respecto al regente (9a), perfectividad (10b), expresa una propiedad o un estado transitorio y el verbo regente tiene carácter evaluativo (11), lleva negación (12), está en subjuntivo (13) o bien refleja un evento repetido (15).

<sup>17</sup> Frente a las relativas que aparecen en construcciones presentativas con *haber*, que sí admiten preposiciones y, por tanto, otras formas del paradigma de los relativos: *Hay autores de los que rara vez se habla* [ABC 01/11/1986].

Estos rasgos peculiares suelen asociarse a una interpretación valorativa o cognitiva del predicado de percepción.

Al retomar las variantes con clítico, en el §4, se comentarán de nuevo estas propiedades, pero antes se presentará el análisis de las cláusulas flexionadas predicativas desde la óptica de su estructura informativa, siguiendo básicamente los postulados de Kleiber (1988) y, sobre todo, de Lambrecht (2000).

### 3. PROPIEDADES INFORMATIVAS

Según Lambrecht (2000), las construcciones estudiadas son a la vez presentativas y predicativas. Por la primera función pragmática sirven para introducir entidades nuevas en el discurso, de ahí que la entidad pivote se realice típicamente como nombre propio o como frase nominal de referente específico. Por la segunda función sirven para proporcionar una información nueva sobre esa entidad. Las cláusulas predicativas presentativas son construcciones singulares porque acumulan dos funciones pragmáticas en una misma estructura argumental, en una única unidad de procesamiento cognitivo. El resultado es una construcción que ofrece ventajas cognitivas, ya que permite asignar doble estatuto al constituyente *pivote*: es objeto de un predicado primario de percepción y sujeto de otro subordinado<sup>18</sup>. Desde el punto de vista de la estructura informativa, la singularidad de las cláusulas flexionadas predicativas reside en que sobre ese constituyente destacado giran dos funciones informativas: tiene el papel de *foco* con respecto al predicado de percepción –es una entidad nueva desde la perspectiva de este predicado– y el de *tópico* con respecto al predicado subordinado. Con palabras de Lambrecht (2000: 51-52):

Au niveau de la structure informationnelle, le référent du SN a une relation de *focus* avec le prédicat principal et une relation de *topique* avec le prédicat subordonné. En tant que focus, le référent joue le rôle de l'entité nouvelle introduite par la proposition principale ; en tant que topique, il joue le rôle de l'entité à propos de laquelle la relative exprime une information nouvelle. La CRP est un exemple particulièrement clair d'une structure formelle qui est *motivée pragmatiquement*.

<sup>18</sup> Esta acumulación de funciones gramaticales no es exclusiva de las cláusulas flexionadas, es compartida por los restantes ejemplos citados de predicaciones secundarias.

## 3.1.

Para entender mejor las propiedades discursivas de las relativas predicativas, Lambrecht (2000) recurre a dos conceptos analíticos desarrollados por Fillmore (1976): el concepto de *mundo del discurso* y el de *contextualización*. Así, los textos reflejan al menos dos mundos discursivos: el mundo en que se produce el texto (cf. el contexto de *enunciación* en la lingüística francesa) y el mundo o los mundos que produce el propio texto (cf. el *enunciado*). Por eso la contextualización puede ser externa, si corresponde a la situación comunicativa, o interna, si viene dada por el universo discursivo creado por el propio texto. Según Lambrecht (2000: 56), en una relativa presentativa se afirma que una entidad cognitivamente inactiva, es decir, una entidad identificable pero que no ha estado activa en el discurso reciente, entra en el campo de percepción (visual o auditivo) del perceptor merced a la actividad en la que está implicado. Observemos el ejemplo de (19), que se ajusta bastante bien al análisis de Lambrecht:

- (18) Habían pasado pocos minutos de las ocho cuando **vi a María que se acercaba**, buscándome en la oscuridad. **Era ya muy tarde para ver su cara, pero reconocí su manera de caminar** [CORDE. 1948. Sábato, Ernesto. *El túnel*. Argentina]

En (18) se dice explícitamente que la oscuridad impedía al perceptor-narrador percibir los rasgos de la cara de María y que fue su peculiar manera de caminar o de aproximarse lo que posibilitó el acto de percepción<sup>19</sup>. Por tanto, la función pragmática de la construcción es concienciar al lector sobre la presencia de una entidad en un mundo interno, el mundo creado en un punto de la narración. De ahí la extrañeza que producirían, con este valor informativo, oraciones con elementos deícticos, como las que contienen marcas que hacen referencia al receptor (*#te veo que te acercas*, *#oyes a María que grita*, cf. Kleiber 1988: §3.2). Examinemos ahora el ejemplo de (19):

<sup>19</sup> Al inicio del capítulo XXVI, en cambio, en el mismo texto de Sábato aparece el siguiente ejemplo:

Cuando nos levantamos de la mesa para caminar por el parque, *vi que María se acercaba a nosotros*, lo que confirmaba mi hipótesis: había esperado ese momento para acercárenos, evitando la absurda conversación en la mesa. [1948. Sábato, Ernesto. *El túnel*. Argentina]

Como se explicará en el §5, el sujeto explícito de la subordinada no está destacado por no constituir un tópico discursivo en esa parte del texto (inicio de capítulo).

- (19) VECINA ¿Y en lo oscuro tú viste las brujas? ¿**Las viste?**  
 CATALINA Sí, cuando se apagó el candil **vi sus bultos corriendo por aquí y las oí que reían y cuchicheaban, pero no sé más. A ninguna le vi la cara.**  
 [CREA. 1978. Miras, Domingo. *Las brujas de Barahona*. España. Teatro]

En este ejemplo la identificación de *las brujas* se realiza pese a que Catalina admite no haberles visto la cara. En su turno de habla confiesa que la identificación se ha basado en los movimientos que vislumbró (cláusula de gerundio predicativo) y en las risas y cuchicheos que percibió (cláusula flexionada predicativa).

La explicación discursiva que estamos adoptando, siguiendo a Lambrecht, permite explicar también el ejemplo de (20), incomprensible tanto en la formulación literal del verbo *morir* ('dejar de vivir') como en las metafóricas ('estar dominado por algo' o 'experimentar violentamente algo'):

- (20) Murió papá, de pura muerte, y yo **lo escuché que se moría...** fue cuando cumplí trece años, y él me miró y después dejó de mirarme... [CREA. 1976. Cabrujas, José Ignacio. *Acto cultural*. Venezuela. Teatro]

Es posible percibir sensorialmente la muerte de alguien por la falta de energía de su mirada o, auditivamente, por sus estertores. Lo que el narrador pretende transmitir en este párrafo es precisamente el impacto que produce que una joven confiese haber percibido sonidos o miradas perdidas que presagian la muerte. Es obvio que aquí la cláusula predicativa no es conmutable por una completiva, ya que *escuché que mi papá se moría* implica que la fuente de información no es el propio hablante, sino terceras personas.

### 3.2.

El carácter presentativo adjudicado a la construcción puede resultar sorprendente si tenemos en cuenta que está admitida la expresión pronominal átona del argumento percibido. Los clíticos se asocian con entidades discursivamente activas, con tópicos ya establecidos. Lambrecht (2000) reconoce lo paradójico de este factor y lo justifica en función de los dos niveles de contextualización que presupone una construcción con relativa predicativa: la entidad presentada es a la vez un tópico establecido

en un mundo discursivo anterior y un objeto de percepción que aparece en otro mundo<sup>20</sup>. Veamos el fragmento de (21):

- (21) —[...] Yo llegué un mes después y **la Lalita** todavía estaba llena de moretones.  
—Te contó que **le** pegué, pero no que **ella** quería matarla a la shapra —dijo Fushía—. Cuando yo me estaba durmiendo, **la vi que agarraba el revólver** y me dio cólera. [CORDE. 1966. Vargas Llosa, *La casa verde*. Perú]

En este fragmento de una novela de Vargas Llosa, una mujer llamada *Lalita* se contextualiza primero como parte de una conversación centrada en el maltrato físico que ha padecido y del que es acusado su marido, Fushía. En su intervención, el marido recontextualiza a la mujer, al culparla de intento de asesinato. Aplicando el análisis de Lambrecht, esa segunda recontextualización sirve para que Fushía intente cambiar la opinión de su interlocutor al inducirlo a visualizar al personaje de Lalita empuñando un revólver.

La explicación discursiva anterior encaja con varios de los ejemplos de corpus examinados, pero no es generalizable a todos los casos documentados. Por una parte, el propio Lambrecht indica que algunos tipos de situaciones no sirven como desencadenantes de una percepción. Según él, predicados mentales como *mentir*, o incluso predicados de actividad física como *sonreír*, *peinarse*, o *endurecer los músculos*, difícilmente permiten advertir la presencia de una persona. Sin embargo, la inserción de este tipo de predicados en esquemas predicativos es frecuente, como se constata en (22):

- (22)a. **El tío Lucho** me acompañó hasta la puerta, me dijo que volviera a almorzar, y cuando, al despedirme, besé a la tía Julia, **lo vi que sonreía** [CREA. 1977. Vargas Llosa, Mario. *La tía Julia y el escribidor*. Perú]  
b. **Nana Mouskouri** se emociona, hay veces que yo **la veo que se emociona en las canciones** [CREA. Oral. España]

<sup>20</sup> En esta línea argumentativa, los ejemplos con el verbo *mirar* que aparecen en acotaciones teatrales reúnen los requisitos textuales señalados por Lambrecht. Los personajes están identificados, porque aparecen en los diálogos de la obra. En las acotaciones, las indicaciones escénicas están dirigidas a directores y actores y hacen referencia a palabras, gestos o movimientos de los actores que representan a los personajes teatrales, lo que implica su recontextualización.

Según Lambrecht (2000: 58) ejemplos como el de (22a) se explican como resultado de la *acomodación* del estatuto cognitivo del referente de una expresión a la función pragmática inherente de la construcción en que aparece. Al producirse una contradicción entre el estatuto cognitivamente activo de una entidad, como ocurre en (22a), y la función presentativa inherente de la construcción, el intérprete se ve forzado a acomodar su valor pragmático al de la construcción, lo que suele crear efectos retóricos específicos<sup>21</sup>, como sorpresa, ironía, o en el caso de (22a) complicidad entre *el tío Lucho* y su sobrino, el narrador, recién casado con su tía (el tío Lucho acaba de presenciar un beso de despedida entre una pareja de recién casados: su hermana y su sobrino). Sin embargo, la cláusula flexionada de (22b) habla de una emoción que el hablante detecta en más de una ocasión y el predicado de percepción indica una constatación intelectual con base sensorial. La acomodación discursiva es parcial, ya que la función presentativa no es aplicable a casos como este.

#### 4. OPCIONES CONSTRUCTIVAS CON CLÍTICO

En este apartado se retoman las combinaciones del tipo *clítico + verbo de percepción + cláusula flexionada*, documentadas principalmente con los verbos *ver* y *oír*<sup>22</sup>. Además de no ajustarse de forma estricta a la interpretación discursiva expuesta en el §3, algunos ejemplos no satisfacen las propiedades sintáctico-semánticas esperables en una predicación secundaria. Se trata de secuencias ya analizadas, como las de (9a, 10b, 11a, 12, 13, 15, 22b), algunas reproducidas de forma abreviada en (23):

<sup>21</sup> Estos efectos también se observan cuando la cláusula de percepción no contiene un clítico. En el fragmento siguiente queda destacada la impresión que produce a un radio-oyente el hecho de oír a su mujer hablando por la radio. Creemos que el efecto de sorpresa quedaría muy atenuado en la versión alternativa: *¡oigo que mi mujer habla con Jesús Álvarez y con usted!*

Soy representante y, el día 5 de junio, iba en mi coche hacia Zaragoza, y a la altura de Alhama de Aragón, voy oyendo la radio, y cuando todo va coincidiendo en su programa con mi calle, casa, piso y... **¡oigo a mi mujer que habla con Jesús Álvarez y con usted!** Y me digo: aquí ha tenido que ocurrir una tragedia, porque a mí mujer la dejé en casa de su madre. [CREA. 1992. Díaz, Lorenzo. *La radio en España* (1923-1993)]

<sup>22</sup> Se documentan menos ejemplos con otros verbos, algunos claros predicados secundarios (i) y otros más parecidos a completivas con prolepsis (ii):

(i) Es más, Filomena, su nana, las escuchó gritarse y bueno, hasta a mí me da pena decírtelo, pero **las escuchó que se gritaban groserías**. [Google. [www.elpisoazul.com/descargas/CUENTOS/Fe.doc](http://www.elpisoazul.com/descargas/CUENTOS/Fe.doc)]

(ii) Una vez en Tokio un hombre se quejaba que **su casa ya no la sentía que era suya**. [CREA. 1996. Sophia. *El Arte de Adivinar con las Cartas*. EE.UU.]



- (23) a. Los amigos míos que se han criado aquí, [...] yo **los veo que han tenido un cambio**.
- b. Pero la señora, yo **la veo que es chévere** y los hijos de ella también, o sea, las personas allá son muy distintas a las del barrio por aquí.
- c. El Minoxidil lo veo que **no** me mejora mucho mi pelo.
- d. Me dio temor: anda demasiado conmigo, **no lo veo que ande con mujeres**, a lo mejor no me salió bien.
- e. **De vez en cuando** oye a Cosme Vila que habla por Radio Moscú.

En (23) el predicado regente no se limita a expresar la captación sensorial directa de un evento que afecta a un participante. El hablante extrae conclusiones personales a partir de su experiencia personal. Las cláusulas flexionadas codifican entidades más abstractas que un simple evento —cf. la anterioridad de (23a), el predicado caracterizador de (23b), la negación de (23c), el subjuntivo y el valor habitual de (23d), la repetición en (23e)— y guardan mayor parecido con los contenidos típicos de una cláusula sustantiva. Pese a las particularidades detectadas, los ejemplos de (23) tienen en común la presencia de un clítico en caso acusativo<sup>23</sup> que posee un antecedente textual y de una cláusula flexionada. Todo parece indicar que en el uso se producen interferencias entre dos esquemas, uno con predicación secundaria y otro con completiva. El resultado es una construcción formalmente equiparable a una cláusula predicativa (*clítico + verbo de percepción + cláusula finita*), pero con un verbo regente con valor epistémico ('darse cuenta', 'opinar', 'juzgar').

Los ejemplos de (23) se distinguen, a su vez, de otra construcción con la que guardan cierto parecido formal, que presenta un clítico neutro sin antecedente textual y que es correferente con la cláusula subordinada. Nos referimos a datos como los de (24):

- (24) a. ¡Pero es ella la primera en ponerse desnuda! En cuanto se van los hijos, se apaga la televisión, y aquí no sé. Cómo

<sup>23</sup> Se registran dativos en dialectos leístas:

<sup>(i)</sup> [...] también yo suelo emplear mucho la palabra, chalao, chalado de las ciencias naturales, borracho de las ciencias naturales y **este muchacho**, pues **le veo que le gusta el campo**, que es capaz de conocer a un animal a distancia... [CREA. Oral. España]

<sup>(ii)</sup> Lo que también recuerdo es que, para intentar centrarme, tranquilizarme, no echar a correr, cuando comencé a hablar ante todo Hollywood, lo único que se me ocurrió para no perderme fue centrar la mirada **el actor Daniel Day-Lewis**, que estaba en la primera fila y siempre ha sido una actor que me gusta mucho. Cuando **le vi que se partía el pecho de risas**, respiré a fondo y me dije: «¡Uf!, esto ha funcionado». [La Vanguardia, 27/03/1994]

dice: “¿Qué pasa?, **¿no lo ves que ya estamos solos?**”

Y ya, claro, ya me incita. Claro claro. [CREA. Oral. España]

- b. –Tú mandas aquí más que tu padre. Para mí por lo menos.  
–Ya. **Ya lo veo que a mi padre lo tiene avasallado.**  
Que ya no me lo deja usted ni cenar, ni puede cerrar el establecimiento, ni marcharse a la cama ni nada. (CORDE. 1956. Sánchez Ferlosio, Rafael. *El Jarama*.)
- c. –Es sólo un reloj -dijo Pepe tímidamente.  
–**Ya lo veo que es un reloj.** ¿Lo trajo usted? (CORDE. 1948-1963. Delibes, Miguel. *La mortaja*. España).

En (24) aparece un clítico neutro catafórico en caso acusativo (*lo*) que concuerda con la cláusula completiva y que no constituye una anticipación o extracción de un constituyente de la subordinada, sino una marca de concordancia del predicado. El clítico no tiene un referente textual al que ligarse y el predicado regente de percepción tiene valor cognitivo, es sinónimo de “darse cuenta, percatarse, comprender”. Con la expresión *ya lo veo* en (24c), el hablante manifiesta su conformidad con una apreciación personal. Por tanto, estamos ante ejemplos claramente diferenciados de los que son estudiados en este trabajo.

Volviendo a la construcción objeto de análisis, los datos indican que se producen cruces entre una cláusula relativa con valor predicativo-presentativo y otra completiva con un constituyente desplazado de su posición, codificado en caso acusativo y convertido formalmente en objeto directo del verbo de percepción (datos de 23). Para justificar este formato a partir de una completiva, en la bibliografía sobre el tema se habla de un caso de *prolepsis* o anticipación del sujeto de la subordinada (cf. Brucart 1999: 432, *NGLE* 2009: §37.4f). El problema que ofrece esta etiquetación es que no suele ir acompañada ni de explicación ni de justificación del fenómeno. En gramáticas normativas el término se ha usado para identificar una variante anómala de otra construcción considerada “normal”. En otro tipo de trabajos, la oración con elemento proléptico se considera derivada de otra básica (se habla de construcciones de elevación de sujeto a objeto, por ejemplo). Por lo que conocemos, sólo se han propuesto explicaciones discursivas para las lenguas clásicas, especialmente para el griego (Panhuys 1984)<sup>24</sup>.

<sup>24</sup> Con un enfoque formal, Maraldi (1986) sostiene que el acusativo proléptico latino no es un esquema derivado, sino una estructura modelada sobre un esquema con acusativo y complemento predicativo. Según Maraldi (1986: 94), “In a way analogous to left-dislocation, the “proleptic accusative” construction contains a

Panhuis (1984) rechaza los análisis tradicionales de la prolepsis como casos de énfasis, relevancia estilística, viveza o incluso prominencia. Y propone una explicación discursiva basada en el “dinamismo comunicativo” y en la progresión tema-remata. Por tanto, para analizar el fenómeno, es preciso traspasar el nivel oracional y situarse en el nivel textual o discursivo. Según este autor, el constituyente temático es aquel que, además de situarse en posición inicial de oración, sirve para conectar el remata con el contexto precedente, con los participantes en la conversación o con elementos de la situación discursiva (lo que otros autores denominan *tópico*). Según Panhuis (1984), cuando la cláusula subordinada contiene un constituyente temático —es un *tópico* discursivo, mencionado previamente en el texto o inferible del contexto verbal o situacional—, posee rasgos discursivos que distorsionan el incremento gradual de rematicidad que debe proporcionar la cláusula subordinada como un todo, lo que favorece que se mueva hacia una posición externa y se sitúe antes del constituyente “nuevo” (el remata, la cláusula subordinada). La anticipación también puede producirse para evitar una lectura errónea de la subordinada. Nos interesa destacar, en particular, una conclusión que se deriva del análisis de Panhuis:

The subject of the subordinate clause occurs thus as an object in the main clause because of its thematic character. If it were rhematic, it would occur in the subordinate clause (Panhuis 1984: 29).

En nuestra opinión, las completivas con constituyente proléptico comparten varios rasgos discursivos con las relativas predicativo-presentativas, lo que puede explicar las interferencias entre las dos construcciones. En el §5 se comprobará que la tesis de Panhuis ofrece puntos de convergencia con el análisis de Lambrecht. Si el sujeto de la predicación secundaria es información nueva, aparece dentro de la cláusula subordinada. Si el sujeto de la predicación secundaria es un *tópico* discursivo, es colocado en primer plano discursivo al recontextualizarlo como sede de la percepción.

---

focused element and a proposition which I take to be “about” this NP, i.e., an open sentence predicated of this NP”. La autora señala también que la estructura es típica de la lengua coloquial y que la elección entre una construcción con o sin prolepsis se debe a motivos pragmáticos.

## 5. ALTERNANCIAS ENTRE CLÁUSULAS

Las cláusulas flexionadas predicativas constituyen una estrategia discursiva que guarda cierto parecido con las estructuras escindidas: en lugar de presentar el contenido proposicional del verbo de percepción de forma global como lo hace una subordinada completiva, la construcción se escinde para destacar a la entidad percibida –Lambrecht le adjudica el papel de *foco* con respecto al verbo de percepción, desde cuya óptica es tratado como si fuese una entidad discursivamente nueva–, y presenta la situación en la que está involucrado como desencadenante de la percepción, por tanto como información nueva sobre la entidad percibida, que desde el punto de vista de la cláusula subordinada es ya un tópico establecido.

Para comprobar las diferencias de tipo informativo, vamos a comparar los dos formatos alternativos: a) clítico con antecedente textual + verbo de percepción + cláusula finita y b) verbo de percepción + cláusula finita. El factor común en (25-26) es el predicado regente *oír* con valor de ‘percibir sensorialmente lo que alguien dice’ y los mismos predicados subordinados –*gritar* en (25) y *llamar* en (26)–.

(25)a. RENATA ¿Tus padres no hacen el amor?

LIDIA No sé, nunca se lo he preguntado. Bueno, una noche **oí que mi padre le gritaba a mi madre: Te mueves menos que un municipal, Charito** [CREA. 1992. Marsillach, Adolfo. *Feliz aniversario*. Teatro]

b. [...] la Charo lo saludaba a él y le sonreía como si fuera a ponerle la banda y la diadema de “Miss Puerto”, y él empezó a mandarle besos con los dedos de la mano, y ella se puso a hacer lo mismo, y así estuvieron un rato, hasta que **oímos a Antonia que gritaba el nombre de Charo y mi nombre**, porque ya era muy tarde y había que irse [CORDE. 1955. Mendicutti, Eduardo. *Fuego de marzo*. España]

(26)a. Todos los días se repetía el mismo ritual: juego, desayuno, siesta en el saúco hasta el atardecer, escalada a los tejados, salto a nuestros hombros, lametones en los ojos, carreras por el jardín, cena y, ¡a la cama!, a descansar en lugar seguro.

Hasta que, una tarde, **oí que Maribel me llamaba con angustia** y, de forma inesperada, cambió todo [CREA. 2001. Pardo de Santayana, José Ignacio. *El beso del chimpancé*]

- b. [...] me levanté, pero no podía contener el sentimiento de estar como ahogada, **entonces oí a Micaela que me llamaba** [CREA. 1986. García Ramis, Magali. *Felices días tío Sergio*. Puerto Rico.]

Las cláusulas flexionadas de (25a) y (26a) presentan la situación percibida de forma holística o integradora: la actividad sonora de una entidad, un grito o una llamada, dirigida hacia otra entidad. Los sujetos de las cláusulas subordinadas, una frase nominal (*mi padre* en 25a) y un nombre propio (*Maribel* en 26a), codifican entidades referenciales que no son tópicos discursivos: en (25a) *mi padre* es mencionado tras una pregunta previa sobre la actividad sexual de los padres de una de las protagonistas y no será retomado en el discurso subsiguiente. En (26a) *Maribel*, la cuidadora de un zoo, está discursivamente inactiva porque en el contexto inmediato se relata la vida cotidiana de dos animales del zoo (dos jinetas) y el discurso subsiguiente tampoco se centra en ella, sino en el suceso que afecta a las dos jinetas protagonistas de ese tramo discursivo. Todos los elementos de la completiva se organizan como un todo, de forma compleja.

En cambio, en (25b) y (26b) se presentan de forma autónoma los emisores del grito y de la llamada, respectivamente, que son entidades cognitivamente reactivadas y a las que el escritor quiere hacer entrar en escena de nuevo porque volverán a tener protagonismo en el discurso subsiguiente. Con respecto al verbo *oír*, las entidades referidas por *Antonia* en (25b) y por *Micaela* en (26b) actúan como puntos de referencia para la percepción y las subordinadas explicitan la actividad sonora detonante de la percepción.

La justificación discursiva de cada una de las estrategias comparadas en el par de (27) es similar:

- (27)a. Desde allí, entre risas y salpicaduras, [Mutarra] llamó a don Alfonso para que se aventurase con él. El infante estuvo dudando un momento, **vio que el prior caminaba ya hacia el puente, de prisa, por la otra orilla** [CREA. 1993. Torbado, Jesús. *El peregrino*. España. Novela]

- b. Fronesis no sabía que después de la muerte de Licario, en un día ciclónico propicio a los excesos, Cemí había conocido en forma definitiva a Ynaca. **Al desembarcar vio a Ynaca que caminaba con la flecha que él había lanzado en la mano.** [CREA. 1977. Lezama Lima, José. *Oppiano Licario*. Cuba. Novela]

El denominador común de (27) es el verbo regente de percepción visual (*vio*) y el predicado subordinado *caminar*, que describe un movimiento. En el fragmento discursivo de (27a) hay dos protagonistas en primer plano – Mutarraff y don Alfonso– y un tercero en segundo plano –el prior–, lo que justifica la visualización de sus movimientos como parte del espacio global complejo que abarca la percepción –la completiva de (27a)–. En cambio, en el fragmento de (27b), el foco se sitúa en el personaje de Ynaca, cuyos movimientos se visualizan. En ese fragmento un personaje denominado Fronesis va en busca de Ynaca, entidad sobre la cual el lector recibe una información desconocida por Fronesis –su relación carnal con Cemí, propiciada por un día ciclónico–. En ese segmento de la narración, Ynaca es traída a primer plano, como cuando se produce un cambio de escenario o de plano cinematográfico, señalado por el adjunto *al desembarcar*.

Por el hecho de contener formas fonéticamente reducidas, el contraste entre los ejemplos del par de (28) es algo diferente. En (28a) el sujeto de la completiva no está explícito, sino que se recupera de las referencias pronominales del contexto previo –es un tópico discursivo–. Ahora bien, aquí no parece justificado usar una estrategia anticipativa –como *la vi que me miraba*–, porque no se requiere una entrada en escena del personaje femenino, cuya proximidad física con respecto al narrador se revela en el fragmento precedente. Por tanto, en (28a) se describe sin más una constatación de origen visual, mientras en (28b) el personaje masculino es recontextualizado al hacerlo entrar en la escena perceptiva mientras se informa de un evento que le afecta y cuyas consecuencias son pertinentes para la continuidad del relato:

- (28)a. **La** tomé de la cintura, **la** atraje hacia mí e intenté besar**la**. No me rechazó pero tampoco me besó: sentí **su** boca fría contra la mía. Al apartarnos, **vi que me miraba sin sonreír**. [CREA. 1977. Vargas Llosa, Mario. *La tía Julia y el escribidor*. Perú].

- b. Otra vez en que volvió a inquirir sobre la edad del niño **lo vi que me miraba receloso**, como cuando él calculaba que yo estaría preñada y no lo estaba: una mirada incrédula, ya dije que recelosa, desconfiada (como si yo pudiera, a menos que fuera bruja, impedir el crecimiento de mi hijo! [CREA. 1993. López Páez, Jorge. *Doña Herlinda y su hijo y otros hijos*. México].

Por tanto, el recurso a la individualización de la entidad percibida marca diferencias en cuanto al estatuto informativo de los referentes y, en cierta medida, también una diferencia en cuanto al tipo de entidad percibida: un evento completo (25a, 26a, 27a, 28a), frente a un individuo que realiza una actividad relevante para la percepción y en buena medida para la evolución del relato (25b, 26b, 27b, 28b).

## 6. CONCLUSIONES

En este trabajo se ha propuesto un análisis discursivo de las cláusulas flexionadas predicativas con verbos de percepción, adaptando una hipótesis de Lambrecht (2000) para el francés. Se ha defendido la tesis de que dicha construcción está pragmáticamente motivada y que comparte propiedades discursivas con los denominados casos de prolepsis. Se ha comprobado también que los datos ofrecen interferencias entre cláusulas flexionadas que actúan como predicaciones secundarias –en contextos de percepción física directa o percepción no epistémica– y completivas de objeto directo –asociadas preferentemente con la percepción epistémica, con base sensorial o puramente cognitiva–. El entrecruzamiento (*blending*) entre ambos esquemas se produce cuando aparece un clítico en acusativo que indica un tópico discursivo.

Interpretadas como estrategias textuales presentativas, las cláusulas predicativas pueden considerarse como un tipo especial de *anagnórisis* en un nivel micro-discursivo: un personaje, a veces el propio narrador, descubre un evento que proporciona información nueva sobre una entidad que debe ser reactivada discursivamente, para lo cual es presentada en primer plano como objeto de percepción. En los casos más claros, el evento expresado en la predicación secundaria es justamente el que permite revelar o confirmar la identidad de dicha entidad; este aspecto es evidente en muchos ejemplos de



corpus, en los que se indica o bien que el perceptor ha entrecerrado los ojos, o bien que está distante –en otra habitación, tras una ventana abierta–, o bien que la percepción tiene lugar en condiciones adversas –en la oscuridad, a través de un resquicio, por una puerta entornada, etc.–. En algunos casos se produce lo que Lambrecht (2000) denomina recontextualización: una misma entidad es tópico establecido con respecto a una secuencia de acontecimientos previa y es objeto de percepción en otro mundo discursivo, en otro escenario; a menudo en los datos aparecen indicaciones temporales que señalan repentinos cambios de perspectiva: *de pronto, de repente, entonces, en ese instante, en ese momento, ya después, instantes después, al levantar la mirada, me doy vuelta y..., giro mi cuerpo con rapidez y..., etc.*

Otro factor destacado es la acomodación del estatuto cognitivo de un referente topical (un clítico) a la función discursiva inherente de una construcción presentativa, lo que crea efectos retóricos concretos, como sorpresa (el contexto menciona sustos, movimientos inesperados, sobresaltos), curiosidad, ironía, complicidad, etc.

Por último, se ha mostrado que una explicación apoyada en el carácter proléptico del sujeto de una completiva requiere el mismo tipo de justificación discursiva que otra basada en una predicación secundaria. La escisión o separación formal del sujeto de la subordinada obedece al carácter de tópico establecido que posee el antecedente como parte del relato previo, a la vez que su referente es presentado como sede o *locus* de la percepción (*Sebastián, que había puesto de nuevo el ojo en la aguja de luz, **los** vio que entraban en la cocina*). La codificación alternativa del evento percibido en forma de completiva implica no sólo una conceptualización holística del percepto, sino también el carácter remático o novedoso del sujeto de la cláusula dependiente (*Primero oí que se abría la puerta y cuando volteé vi que entraban Santana y Majorro seguidos de varios policías*).





## CORPUS

- BDS: Base de datos sintácticos del español actual. Universidade de Santiago de Compostela  
<[www.bds.usc.es](http://www.bds.usc.es)>.
- CORDE: REAL ACADEMIA ESPAÑOLA: Banco de datos (CORDE) [en línea]. *Corpus diacrónico del español*.  
<<http://www.rae.es>> [junio-julio de 2013].
- CREA: REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual*.  
<<http://www.rae.es>> [junio-julio de 2013].



## REFERENCES

- Auwers, Johan van der (1985): "The predicative relatives of French perception verbs", en A. Machtelt Bolkestein, C. de (Casper) Groot y J. Lachlan Mackenzie (eds.): *Predicates and Terms in Functional Grammar*. Dordrecht: Foris, pp. 219-234.
- (1993): "Les relatives prédicatives du français", en Claude Muller y Daniel Roulland (eds.): *Subordinations*. Rennes : Presses Univ. Rennes et CERLICO, pp. 265-277.
- Brucart, Josep M<sup>a</sup> (1999): "La estructura del sintagma nominal: las oraciones de relativo", en Ignacio Bosque y Violeta Demonte (eds.): *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe. Cap. 7, pp. 395-522.
- Carrasco Gutiérrez, Ángeles (2011a): «*Ver* como verbo de percepción epistémica primaria», en Pablo Cano López *et al.* (eds.): *Actas del XXXIX Simposio Internacional de la SEL, Santiago de Compostela, 1-4 febrero de 2010*. Edición electrónica en CD. Santiago de Compostela: Unidixital.
- (2011b): "La percepción de estados", en María V. Escandell Vidal *et al.* (eds.): *60 problemas de gramática dedicados a Ignacio Bosque*, Madrid: Akal, pp. 198-204.
- De Vries, Mark (2002): "Patterns of relative clauses". *Linguistics in the Netherlands* 18, pp. 231-243.  
En <<http://odur.let.rug.nl/dvries/pdf/2001-patterns-rc-lin.pdf>>
- Fernández Jaén, Jorge (2012): *Semántica cognitiva diacrónica de los verbos de percepción física del español*. Tesis Doctoral. Univ. de Alicante.

- Fillmore, Charles J. (1976): "Pragmatics and the description of discourse", en S. Schmidt (ed.), *Pragmatik II*. Munich: Wilhelms Fink Verlag. Reprinted in P. Cole (ed.) (1981) *Radical Pragmatics*. New York: Academic Press.
- Hanegreefs, Hilde (2008): *Los verbos de percepción visual. Un análisis de corpus en un marco cognitivo*. Tesis doctoral. KULeuven.
- Kleiber, Georges (1988): "Sur les relatives du type "Je le vois qui arrive", *Travaux de Linguistique* 17, pp. 89-115.
- Lambrecht, M. Knud (2000): "Prédication seconde et structure informationnelle: la relative de perception comme construction présentative", *Langue française* 127, pp. 49-66.
- Maraldi, Mirka (1980): "The complement structure of perception verbs in Latin", en Gualtiero Calboli (ed.): *Papers on Grammar I*, Bologna: Editrice CLUEB, 47-79.
- (1986): "The proleptic accusative: problems of structural analysis", Gualtiero Calboli (ed.): *Papers on Grammar II*. Bologna: Editrice CLUEB, 87-105.
- NGLE = RAE y Asociación de Academias de la lengua española (2009): *Nueva gramática de la lengua española*. Madrid: Espasa Libros.
- Panhuis, Dirk (1984): "Prolepsis in Greek as a Discourse Strategy", *Glotta* LXII, pp. 26-39.
- Pompei, Anna (2009): "Relative clauses", en Philip Baldi y Pierluigi Cuzzolin (eds.): *New perspectives on historical Latin syntax*. Berlin, New York: Mouton de Gruyter, Vol. 4, pp. 427-535.
- Rafel, Joan (1999): "La construcción pseudo-relativa en romance", *Verba* 26, pp. 165-192.
- Silva-Corvalán (1999): "«Copias pronominales» en cláusulas relativas en español conversacional de Santiago de Chile", en José Antonio Samper Padilla y Magnolia Troya Déniz (coords.): *Actas del XI Congreso Internacional de la ALFAL. Las Palmas de Gran Canaria, 22-27 de julio de 1996*, Las Palmas de Gran Canaria: Servicio de Publ. de la UPGC. Tomo 1, pp. 447-454.
- Tekavčić, Pavao (1980): *Grammatica storica dell'italiano. II. Morfosintassi*. Bologna: Società editrice il Mulino. 2<sup>a</sup> ed. reformada.

## Sobre las formas en *-ra* en el español de Galicia<sup>1</sup>

Guillermo Rojo  
Victoria Vázquez Rozas  
*Universidade de Santiago de Compostela*

**Abstract:** The article analyzes the frequency, sociolinguistic distribution and values of the *-ra* verb forms in a sample of spoken Galician Spanish made up of 18 semi-directed interviews from the PRESEGAL–PRESEEA corpus. A first quantitative analysis of the data shows that the frequency of the etymological indicative *-ra*, a value which has survived both in Galician and in Spanish spoken in Galicia, varies considerably from one speaker to another. Furthermore, individual preferences neither seem to be related to age, gender or educational level of the speakers, nor –apparently– to how frequently they use the subjunctive *-se* form, which on the whole is more common in this variety of Spanish than in others. Nevertheless, by bringing the *había* + participle data into the picture we can make sense of the actual distribution and values of both *-ra* and *-se* verb forms in the corpus.

**Keywords:** Spanish from Galicia, verb, indicative, subjunctive, pluperfect.

---

<sup>1</sup> Estudio perteneciente al proyecto ESLORA, financiado por el Ministerio de Economía y Competitividad (FFI2010-17417).



## 1. LAS FORMAS EN -RA

Como se ha señalado repetidamente, uno de los procesos más llamativos e interesantes en el verbo románico es aquel mediante el cual la forma *amara*, descendiente del llamado tradicionalmente pluscuamperfecto de indicativo latino (*amaveram*), evoluciona en español hasta adquirir los valores temporales y modales correspondientes al pretérito de subjuntivo, con lo que entra en conflicto con la forma *amase*, procedente, por su parte, del pluscuamperfecto de subjuntivo latino (*amavissem*). Reducido a sus aspectos más visibles, se da, en primer lugar, el cambio a un valor temporal básico de anterioridad al origen desde un valor de anterioridad a una referencia anterior al origen; en segundo lugar, y más importante, la forma va adquiriendo, en una evolución cuyos detalles no conocemos todavía por completo, los valores modales típicos del modo subjuntivo, sin que ello implique la desaparición completa de los valores indicativos<sup>2</sup>.

Con los datos de los que disponemos, en la actualidad las formas en *-ra* (simple y compleja) son claramente mayoritarias en todas las variedades del español, hasta el punto de que son muchos los textos en los que suponen el 100% de los usos posibles para ambas. Naturalmente, como sucede en todos los procesos de cambio, hay diferencias, que han sido relacionadas con diversos factores. Se ha apuntado repetidamente que el descenso de *-se* resulta más acusado en América que en España, pero, sin negar lo anterior, parece más importante la relación con el grado de formalidad de los textos, de modo que en todos los países hispánicos los porcentajes más altos de *-ra* se dan en la lengua oral (cf. Rojo 2008: 176-177). Por citar un caso muy llamativo, la *Base de datos sintácticos del español actual* (BDS) muestra que la forma *amase* está completamente ausente en el libro de relatos *Querido Diego*, de Elena Poniatowska, publicado en 1987, mientras que supone el 92,31% de los usos posible para *amase* y *amara* en *Tiempo nublado*, un texto ensayístico de Octavio Paz, publicado en 1984.

Evidentemente, que dos textos mexicanos, muy próximos en el tiempo, presenten tal divergencia, solo puede explicarse a través de la actuación de factores relacionados con el tipo de texto, el tipo de lengua (cf. Rojo, 1996: 685). La forma en *-se*, muchísimo menos habitual, es sentida probablemente como más elegante, más propia de la lengua culta.

---

<sup>2</sup> Para una visión general y sintética de los cambios producidos en la evolución del sistema verbal latino a los correspondientes a español por un lado y gallego por otro, vid. Veiga (2011).



Tal como se manifiesta en español, las formas en *-ra* se van acercando a usos y valores propios del subjuntivo a partir de su utilización como formas indicativas dislocadas, con lo que alteran su valor temporal y le añaden un factor modal de irrealidad (o, si se prefiere, no-realidad), como sucede típicamente en las oraciones condicionales (cf. Rojo y Montero 1983). Una indicación clara de la medida en la que las formas en *-ra* van incrementando su uso como formas típicamente subjuntivas y, consiguientemente, provocando la reducción de las formas en *-se* puede obtenerse de los datos sobre la combinación *para que + -ra / -se*, tal como propone Becker (2008: 148). En el cuadro 1 figuran los datos correspondientes a estas combinaciones a lo largo del tiempo que hemos obtenido del *Corpus del español* (CE)<sup>3</sup>.

	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
<i>-se</i>	100	96,6	99,6	95,16	88,54	96,28	66,34	9,65
<i>-ra</i>	0	3,4	0,4	4,84	11,46	3,72	33,66	90,35
N =	31	65	260	1755	1099	1020	1426	943

**Cuadro 1.** Porcentajes correspondientes a la expresión *para que + -se/-ra* en diversos períodos del español.

Fuente: *Corpus del español* (comprobado el 23/7/2013). Elaboración propia.

Como se puede observar, en esta expresión, que rige subjuntivo, el porcentaje correspondiente a las formas en *-ra*, que alcanza un máximo del 11,46% hasta el siglo XVIII, supone el 90% en el conjunto del siglo XX<sup>4</sup>. Es, con las reservas mencionadas en la nota 4, representativo de la distribución general que presentan estas dos formas a lo largo de la historia de la lengua.

Así pues, en líneas generales el proceso consiste en que las formas en *-ra* pierden su valor indicativo inicial y adquieren valores subjuntivos. Como ha señalado Veiga (entre otros, 1996, 2006), no tiene sentido considerar que la adquisición de los valores subjuntivos se produce como consecuencia de la pérdida de los indicativos. La coexistencia de ambos bloques de valores es perfectamente posible, como, según veremos, muestran el gallego, el español de Galicia y la propia historia del español.

<sup>3</sup> Aunque el perfil general es el mismo, los recuentos que hemos obtenido difieren de los que proporciona Becker. Sus datos arrojan un total de 2959 casos y los que aparecen en el cuadro número 1 suman 6599. Probablemente se debe a que proceden de opciones distintas de la aplicación de consulta del *Corpus del español*. Pero la línea evolutiva es, como se ha dicho, básicamente la misma.

<sup>4</sup> El *Corpus del español* solo permite búsquedas selectivas por siglos, lo cual puede ocultar importantes diferencias en las características de los fenómenos estudiados.

En español actual, pues, la forma en *-ra* posee básicamente el valor del llamado pretérito de subjuntivo, lo mismo que la forma en *-se*. Pero existen algunos factores adicionales que es necesario tener en cuenta. En primer lugar, la forma en *-ra* ha conservado algunos valores indicativos, aunque dislocados (cf. Rojo 1974, Rojo y Veiga 1999), en los verbos modales: *debiera / pudiera / quisiera trabajar más*, usos en los que alterna con *debería / podría / querría*, pero no con *debiese / pudiese / quisiese*, y algunos otros en ciertas zonas de América (cf. Nowikov 1991, De Sterck 2000, NGLE: 24.2). En segundo lugar, como consecuencia de un proceso que comenzó en el siglo XIX<sup>5</sup>, ciertos registros lingüísticos –inicialmente lengua literaria, pero actualmente más bien lenguaje periodístico– muestran casos de recuperación de los antiguos valores de pluscuamperfecto de indicativo (o de pretérito, cf. infra) en expresiones del tipo *Reseñaron el libro que publicara en 1980* y similares. Por último, la incompreensión de estos últimos usos –lógica en hablantes que no pueden captar el aroma arcaizante de estas construcciones– ha llevado a la aparición de formaciones del tipo *Reseñaron el libro que publicase en 1980*<sup>6</sup>.

Estos elementos adicionales, que resultan del mayor interés teórico, no pueden hacernos olvidar las características básicas de este proceso en español: las formas en *-ra* han perdido los valores indicativos en la lengua habitual (salvo los correspondientes a los verbos modales) y han ocupado casi por completo el ámbito de las formas en *-se* y lo han hecho de un modo que se relaciona con la variabilidad diatópica, pero también, y probablemente sobre todo, la diastrática y diafásica. Siempre en términos generales, las formas

<sup>5</sup> En la misma línea de los ocho casos localizados por Girón Alconchel (2000) en textos hispanoamericanos de los siglos XVII y XVIII, González Ollé (2012) muestra con claridad que es posible encontrar casos de *cantara* con valor de pluscuamperfecto de indicativo en todas las épocas de la lengua, por lo que, en su opinión, sería mejor sustituir las habituales alusiones a «pérdida» y «resurrección» o «recuperación» de estos usos (a partir del XIX) por los de «decrecimiento» y «crecimiento» o «revitalización», respectivamente. La documentación que aporta es muy amplia y no deja lugar a dudas acerca de la pervivencia de estos usos en textos que se sitúan a lo largo de todo el período comprendido entre finales del XV y comienzos del XIX, sobre todo “en tres tradiciones discursivas: el *Romancero*, la *Historia* y el *Derecho*” (p. 131). Sin embargo, la existencia de empleos de *cantara* con valor etimológico en textos pertenecientes a tradiciones discursivas marcadas por un cierto aire arcaizante no anula por completo el carácter especial que poseen estos usos en los registros a los que se alude en el texto. Están en la misma línea, aunque ahora con un cambio notable tanto en la frecuencia como en los registros en que se pueden documentar. Por otra parte, González Ollé detecta usos de *cantase* con valor indicativo en textos bastante más antiguos que los que manejábamos hasta ahora, lo cual refuerza, a nuestro modo de ver, el carácter especial de estos empleos (cf. nota 6).

<sup>6</sup> Sin duda más frecuentes en los últimos años, pero localizables desde hace bastante tiempo. Cf., entre otros, Bejarano (1962), Hermerén (1992), Rojo (2011b) y ahora González Ollé (2012: 121 y sigs.).

en *-se* se consideran más cultas y elegantes, lo cual explica la resistencia que muestran en algunos registros e incluso su invasión del antiguo territorio indicativo de las formas en *-ra*.

La evolución experimentada por el español es radical por ambos extremos (pérdida casi total de valores indicativos y conquista de los usos inicialmente reservados a las formas en *-se*), como muestra lo que se puede observar en romances próximos. El portugués conserva las formas en *-ra* como *mais-que-perfeito de indicativo*, esto es, con valor temporal de situación anterior a una referencia anterior al origen, pero su uso es minoritario y literario frente al llamado *mais-que-perfeito composto de indicativo* (*tinha/havia amado*) (cf. Cunha / Cintra 1984: 455 y sigs.). La lengua literaria ofrece también ejemplos de usos modalizados de *amara*, en los que se hace equivalente al “futuro do pretérito de indicativo e do pretérito de subjuntivo, o que serve hoje como traço estilístico de linguagem solene” (Bechara 1999: 279). Sin embargo, esos usos han pasado a la lengua habitual en expresiones fijas del tipo *Quem me dera!* (= *Quem me desse!*, Cunha / Cintra 1984: 456). En palabras de Becker (2008: 166), la forma en *-ra* “was not only blocked from evolving into a general marker of non veridicality, but also receded during the 19<sup>th</sup> century from conditional contexts” y su uso con valor exclusivamente temporal está “relegated to formal written registers”.

En catalán las formas en *-ra* han sido sustituidas por las compuestas (*havia* + participio) en los valores indicativos y las formas en *-se* son las únicas que aparecen con el valor de pretérito de subjuntivo. La *Gramàtica del català contemporani* (= Solà / Lloret / Mascaró / Pérez-Saldanya 2002, I: 636) señala, de todas formas, que, en la variedad valenciana, el imperfecto de subjuntivo “presenta dues solucions: *ra/se*” de ámbito general.

El gallego muestra en este punto la situación más amplia posible. De una parte, conserva con plena vigencia los valores propios del antecopretérito de indicativo, esto es, la expresión de una situación anterior a una referencia anterior al origen (*Cando chegamos, eles (xa) marcharan*). Asociado a este valor, lo mismo que ocurría ya en latín<sup>7</sup>, estas formas pueden presentar usos con un valor temporal de anterioridad al origen, es decir, equivalentes a grandes rasgos a los del pretérito de indicativo: *Lémbrome dunha vez que*

<sup>7</sup> Según Bassols (1956: § 327), el pluscuamperfecto “expresa generalmente acciones pasadas y anteriores a otra acción o determinación; a veces, sin embargo se desvanece este último significado”. Esto es, no presenta carácter relativo y, por tanto, sus usos “pueden equivaler a un imperfecto o perfecto” (ibídem, § 329).

*fóramos a Montelouro* (Álvarez / Regueira / Monteagudo 1995: 370). De estos usos se ha dicho que producen un “distanciamiento psicológico” (Álvarez / Regueira / Monteagudo 1995: 370), poseen un valor “evocativo” (Álvarez / Xove 2002: 300) o son propios de un grado más bajo de focalización discursiva (Pollán 2001, 2003).

Al lado de estos valores rectos, aparecen los también indicativos, pero con una dislocación que les incorpora valores modales de no-realidad en verbos modales (*Quixera deixar este traballo; Deberan ter algo máis de paciencia; Puidera non ser así*) (cf. Álvarez / Regueira / Monteagudo 1995: 370; Álvarez / Xove 2002: 300). Lo mismo que sucede en español, en estos usos se da la alternancia con las formas *amaría* y *amaba*, pero no con *amase*: *Quixera ~ querería ~ quería ~ \*quixese deixar este traballo*<sup>8</sup>.

En un proceso paralelo al que tuvo lugar en español y cuando menos se inició en portugués y en algunas variedades del catalán, la forma *-ra* del gallego comenzó a ser utilizada en empleos dislocados en estructuras como las condicionales o las concesivas y de ahí pasó a confluír con los empleos propios del imperfecto de subjuntivo<sup>9</sup>:

(1)

(a) *Se tivera ~ tivese posibilidade de facelo, voltaría á miña casa.*

(b) *Oxalá viñese ~ viñera axiña!*

(c) *Dixéronme que esperase ~ esperara a súa volta.*

Esta alternancia ha sido considerada efecto de la influencia del español y consiguientemente rechazada en las orientaciones normativas<sup>10</sup>, pero es plenamente aceptada en muchos tratados gramaticales (cf. Álvarez / Regueira

<sup>8</sup> El uso de otros verbos en estas estructuras, como en los ejemplos del tipo *Fora ~ sería ~ era mellor pasar antes polo zapateiro, pero agora xa non hai que facer*, recogidos en Álvarez / Xove (2002: 300) resulta claramente arcaizante. Cf. Xove (1986: 142 y sigs.) para los que aparecen en la obra de Rosalía de Castro.

<sup>9</sup> Dada la estructura del verbo gallego y la ausencia de formas compuestas, la actuación de la dislocación sobre las formas en *-ra* y en *-se* hace que puedan servir para expresar condicionales o concesivas tanto potenciales e irreales de presente o futuro como irreales de pasado, esto es, contrarias a los hechos. Son equivalentes en español de (1a) tanto *Si tuviera esa posibilidade...volvería* como *Si hubiera tenido esa posibilidade... habría vuelto*. Por la misma razón, (1b) puede equivaler a *¡Ojalá viniera pronto!* y a *¡Ojalá hubiera venido pronto!* Por su parte, la conservación plena del valor de pluscuamperfecto de indicativo hace que *Dixéronme que saíras* tenga como posibles equivalentes en español *Me dijeron que salieras ~ salieses* y *Me dijeron que habías salido*. Lo mismo sucede en español de Galicia (cf. Rojo 2004: 1095, nota 19).

<sup>10</sup> Las *Normas ortográficas e morfolóxicas do idioma galego* (RAG / ILG 2003: 111), aunque reconocen que “[a]s formas en *-ra* funcionan tamén como pretérito de subxuntivo”, recomiendan “manter ben discriminados os usos dos dous tempos (*-ra* coma antepretérito de indicativo e *-se* coma pretérito de subxuntivo)”.

/ Monteagudo 1995: 317; Álvarez / Xove 2002: 301)<sup>11</sup>. Parece claro que resulta perfectamente explicable como un desarrollo propio, que tiene plena justificación histórica y paralelos en otras lenguas románicas (cf. Veiga 2005: 33-36). Los datos del gallego actual confirman plenamente este enfoque. En la versión del CORGA (*Corpus de referencia do galego actual*) anotada y desambiguada, el valor de pretérito de subjuntivo aparece en la forma *-se* en el 82,36% de los casos y en la forma *-ra* en el 17,64% restante. Teniendo en cuenta que esta versión contiene únicamente textos escritos, la convivencia de las dos formas y el predominio de las originariamente subjuntivas resulta perfectamente comprensible.

Valor de irrealidad en un punto anterior al origen presentan también empleos del tipo *Estudiaras cando tiñas tempo!*, equivalentes al español *¡Haber estudiado cuando tenías tiempo!*

La organización de la estructura y uso de las formas verbales del español de Galicia está basada inicialmente en el gallego (cf. Rojo 2004). Por tanto, no es de extrañar que las formas en *-ra* presenten en esta variedad los usos que corresponden a las formas morfológicamente equivalentes del verbo gallego y que hemos esquematizado en párrafos anteriores, aunque, como es lógico, sometida a ciertos reajustes. En efecto, a ese proceso básico hay que añadir algunos factores adicionales que es forzoso tener en cuenta. De un lado, está bastante extendida entre los hablantes la idea de que usos como *Cuando llegamos ellos ya se fueran* o *Me enteré de que vinieras a verme* son ‘incorrectos’, propios de un español contagiado de rasgos galleguizantes, y que deben ser sustituidos por *Cuando llegamos ellos ya se habían ido* o *Me enteré de que habías venido a verme*. Como consecuencia de ello, lo esperable es que podamos documentar diferencias en la frecuencia de estos usos que pueda ser correlacionada con factores de carácter sociocultural: a mayor conocimiento del español estándar (de España), menor frecuencia de formas en *-ra* con valor indicativo y mayor empleo de las formas del tipo *había* + participio en su lugar. Pero al tiempo, en un curioso movimiento circular, la idea de que *amara* debe ser sustituida por *había amado* lleva, en algunos casos, a utilizar esta última forma en contextos que, si bien se corresponden con usos de *amara* en gallego, tienen, en cambio, a *amé* o *amaba* como equivalente en español

<sup>11</sup> Pero no todos. Freixeiro (2000: 358) indica que “[t]anto na fala como na lingua literaria as formas de antepretérito (-ra) e de pretérito de Subx. (-se) conflúen, por interferencia do español, nos valores deste último (“raísmo” ou uso de -ra por -se), coincidencia que debe ser totalmente rexeitada desde o modelo de lingua culta que se está a conformar”.

estándar, como se observa en, por ejemplo, *¿Te acuerdas de aquella señora que habíamos conocido cuando habíamos ido a Segovia?* (cf. Blas 1991 y 1993, Rojo 2004: 1096).

Son, por tanto, varios y muy diversos los fenómenos susceptibles de enfoque sociolingüístico que el español de Galicia presenta en los usos de las formas en *-ra*. Su análisis completo requeriría estudiar, por lo menos, los aspectos mencionados en (2):

(2)

- (a) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas *amara* y *amase* con valor de pretérito de subjuntivo.
- (b) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas de tipo *amara* con valor temporal de anterioridad a una referencia anterior al origen en comparación con las correspondientes a las formas del tipo *había amado*.
- (c) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas de tipo *amara* con valor temporal de anterioridad al origen.
- (d) Frecuencia general y distribución por sexo y, edad y nivel sociocultural de las formas de tipo *había amado* con valor temporal de anterioridad al origen (~ c).
- (e) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas en *-ra* correspondientes a los verbos modales en comparación con las formas en *-ría* de esos mismos verbos.
- (f) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas en *-ra* y en *-se* en el condicionante de las condicionales y en el antecedente de las concesivas irreales.
- (g) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas compuestas en *-ra* y en *-se* en el condicionante de las condicionales y el antecedentes de las concesivas irreales.
- (h) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas en *-ra* y en *-se* de las cláusulas incrustadas con dependencia de predicados regentes con valores de anterioridad al origen (*Nos pidió que la perdonásemos ~ perdonáramos*).

- (i) Frecuencia general y distribución por sexo, edad y nivel sociocultural de las formas (simples y compuestas) en *-ra* y en *-se* en usos de subjuntivo regido no directamente relacionados, como *h*, con casos de *consecutio temporum*.

Dadas las limitaciones de espacio de este trabajo, nos limitaremos aquí a analizar la forma en que se manifiestan los puntos a-d en las 18 entrevistas correspondientes al nivel sociocultural alto de la ciudad de Santiago de Compostela (PRESEGal), integradas en el *Proyecto para el estudio del español de España y América* (PRESEEA)<sup>12</sup>. Cuando no se indique explícitamente lo contrario, los datos y explicaciones sobre las formas *-ra* y *-se* en los apartados siguientes se refieren exclusivamente a las formas simples.

## 2. FRECUENCIA DE LAS FORMAS EN *-RA* EN ESPAÑOL

La situación de los datos que hemos obtenido de las entrevistas correspondientes a los hablantes con estudios universitarios del español de Galicia requiere su comparación con lo que podemos observar en el español general tal como se presenta en la actualidad. El *Corpus del español* (CE), parcialmente lematizado, nos permite obtener una visión general de la distribución de las formas en *-ra* y en *-se* en el siglo xx y también la correspondiente a distintos tipos de texto. Los datos son los que aparecen en el cuadro 2:

	<i>General</i>	<i>Académico</i>	<i>Prensa</i>	<i>Ficción</i>	<i>Oral</i>
<i>-ra</i>	89,39	92,43	88,64	88,04	90,98
<i>-se</i>	10,6	7,56	11,35	11,95	9,01
N =	42 195	5209	7020	21 167	8349

**Cuadro 2.** Porcentajes correspondientes a las formas en *-ra* y *-se* en textos del siglo xx.

Fuente: *Corpus del español*. Elaboración propia. Datos obtenidos el 26/7/2013

Es evidente que fundir datos correspondientes a un período de cien años resulta excesivamente amplio, pero el CE no admite consultas más refinadas en este sentido y tampoco facilita la selección por países. Así

<sup>12</sup> Vid. infra, apdo. 3 la descripción de la muestra utilizada.



pues, lo que tenemos es que la forma en *-ra* muestra un dominio muy marcado, que se da tanto en general como en cada uno de los tipos de texto diferenciados. Incluso en la línea de una perspectiva muy amplia, en la que, por su propia generalidad, los rasgos específicos tienden a desdibujarse, no deja de sorprender la proximidad de los porcentajes de uso en los distintos tipos textuales, así como que sean los textos académicos los que presenten el porcentaje más alto de las formas en *-ra*, seguidos por los textos orales<sup>13</sup>.

El *Corpus de referencia del español* actual (CREA) permite seleccionar por países y tipos de texto, pero la versión pública no está lematizada todavía, de modo que hay que usar estrategias alternativas, como la utilizada en Rojo (2008), que consiste en hacer búsquedas para un grupo de formas específicas de verbos de frecuencia alta. Así, los datos correspondientes a los once verbos estudiados en Rojo (2008) en los textos del CREA son los que aparecen en el cuadro 3:

	<i>Ficción</i>	<i>Libros (no ficción)</i>	<i>Periódicos</i>	<i>Revistas</i>	<i>Orales</i>
<i>-ra</i>	81,49	83,52	85,43	89,68	90,39
<i>-se</i>	18,51	16,48	14,57	10,32	9,61
N =	41 296	26 096	10 593	2499	2507

**Cuadro 3.** Porcentajes correspondientes a las formas de primera y tercera singular y tercera de plural en *-ra* y *-se* de los once verbos de frecuencia alta en la totalidad del CREA analizados en Rojo (2008). Fuente: CREA. Elaboración propia.

Se observa que los porcentajes de las formas en *-se* son siempre superiores a los que se dan en el *Corpus del español* (para todo el siglo xx y de todos los países hispánicos) y también que, de acuerdo con lo esperado, son los textos orales los que muestran cifras más altas para las formas en *-ra*, con algo más de un 90% del total.

Las estadísticas procedentes de la BDS, cuyos textos corresponden a los períodos centrales del CREA, dibujan un panorama bastante parecido para los textos españoles, que es el que aparece en el cuadro 4.

<sup>13</sup> El análisis de los datos del *Corpus diacrónico del español* (CORDE) en tramos temporales más reducidos muestra fuertes diferencias en el interior del siglo xx. Las formas en *-ra* pasan de un 51% en el primer cuarto de siglo al 82,31% en el último (cf. Rojo 2008: 169, tabla 4). Por otro lado, la consideración de qué son textos orales en el CE resulta más que discutible (cf. Rojo 2010).



	<i>General</i>	<i>Ensayo</i>	<i>Narración</i>	<i>Prensa</i>	<i>Teatro</i>	<i>Oral</i>
<i>-ra</i>	79,74	62,37	87,58	61,07	77,22	87,37
<i>-se</i>	20,26	37,63	12,42	38,93	22,78	12,63
N =	2493	497	1337	131	338	190

**Cuadro 4.** Porcentajes correspondientes a las formas en *-ra* y *-se* en los textos españoles de la BDS. Fuente: BDS. Elaboración propia.

Hay, sin duda, diferencias entre los datos del cuadro 4 (todos los casos de textos españoles de un corpus muy pequeño -algo menos de un millón y medio de formas) y los del cuadro 3 (algunas formas de un subconjunto de verbos muy frecuentes de textos de todos los países hispánicos en un corpus de algo más de 150 millones de formas), pero el dato que aquí nos interesa fundamentalmente, el que corresponde a la frecuencia de las formas en *-ra* en textos orales resulta sorprendentemente próximo. Así pues, el telón de fondo sobre el que vamos a proyectar los datos de PRESEGAL muestra cifras comprendidas entre el 90,98% del CE para textos de todos los países hispánicos y todo el siglo xx y el 87,37 % que encontramos en la parte oral del componente español de la BDS. Son cifras realmente congruentes a pesar de la considerable diferencia que existe entre los textos procesados en cada uno de estos conjuntos textuales. Próximos a ellos están también los datos que conocemos de estudios cuantitativos realizados sobre este tema en otros corpus orales del español contemporáneo (cf., entre otros, de Sterck 2000).

### **3. FORMAS EN -RA EN EL ESPAÑOL DE GALICIA (NIVEL SOCIOCULTURAL ALTO EN PRESEGAL)**

#### **3.1. La muestra**

Los datos aquí analizados pertenecen al corpus PRESEGAL (*Proyecto para el estudio sociolingüístico del español de Galicia*), constituido por 54 entrevistas semidirigidas realizadas a hablantes de español de Santiago de Compostela. Siguiendo la estructura prevista en el proyecto PRESEEA, en el que se integra, los informantes se reparten de forma homogénea en tres niveles de estudios -primarios, medios y universitarios-, en tres grupos de edad -de 20 a 34 años, de 35 a 64 y de 65 en adelante- y entre mujeres y hombres.

La muestra utilizada para el presente estudio comprende las 18 entrevistas de hablantes universitarios, 9 mujeres y 9 hombres, distribuidos en igual número en los tres grupos de edad indicados. Para facilitar el análisis de los ejemplos citados en el texto, incluimos en apéndice la distribución de las entrevistas y sus denominaciones. Las intervenciones de los informantes en las 18 entrevistas suponen un total de algo más de 220 000 formas ortográficas, que han sido etiquetadas morfosintácticamente mediante el programa FreeLing (cf. Padró y Stanilovsky, 2012), adaptadas posteriormente, mediante unas rutinas desarrolladas por Paula Santalla, al etiquetario que se utiliza habitualmente en nuestro grupo de investigación (cf. Santalla y Sotelo 2006) y revisadas manualmente en los casos implicados en este estudio. Como resultado de todo ello, hemos trabajado sobre un corpus que tiene un total de 167 000 unidades lingüísticas (excluyendo signos de puntuación, cifras, etc.).

### 3.2. Perspectiva general

Los datos generales que hemos obtenido en las 18 entrevistas correspondientes a hablantes con estudios universitarios realizadas en Santiago de Compostela para el proyecto PRESEGAL son los que aparecen en el cuadro 5, en el que incluimos también los datos correspondientes a otras formas relacionadas.

Formas en <i>-ra</i> (simples)	220
Formas en <i>-se</i> (simples)	72
Formas <i>hubiera</i> + participio	26
Formas <i>hubiese</i> + participio	27
Formas <i>había</i> + participio	179
Total	524

**Cuadro 5.** Frecuencias de diferentes tipos de formas en el nivel alto de PRESEGAL (Santiago de Compostela). Elaboración propia.

En lo que respecta a la distribución de usos entre las formas en *-ra* y las formas en *-se*, se observa que las formas en *-ra* presentan un porcentaje de empleo inferior al que podemos ver en el componente oral de otras variedades del español: un 73,34%, a pesar de que en los recuentos globales entran también los usos plenamente indicativos que no se han conservado o lo han hecho en mucho menor medida en otras variedades.

Pero la distribución en el interior de la muestra dista mucho de ser homogénea. Si atendemos a la variable sexo, los porcentajes de uso son los que aparecen en el cuadro 6:

	<i>Hombres</i>	<i>Mujeres</i>
<i>-ra</i>	67,23	80,93
<i>-se</i>	32,77	19,07
N =	119	173

**Cuadro 6.** Porcentajes correspondientes a las formas en *-ra* y *-se* por sexos en PRESEGAL. Elaboración propia.

La diferencia es bastante fuerte y, además, se manifiesta en varios aspectos. La distancia en la distribución de las formas se acerca a los 14 puntos porcentuales entre los dos sexos. Pero se observa además que hay bastante distancia en el número total de formas que corresponde a cada uno: 119 frente a 173 formas simples en *-ra* o *-se*. Es cierto que la suma de unidades lingüísticas correspondientes a los informantes es mayor en las mujeres (88 655) que en los hombres (78 748), pero las 10 000 unidades de distancia no son suficientes para explicar las diferencias, de modo que parece necesario concluir provisionalmente que las mujeres emplean bastante más estas dos formas y, además, prefieren la forma *-ra* en mayor medida que los hombres.

También se observan diferencias importantes entre los tres grupos de edad considerados, como muestra el cuadro 7:

	<i>Grupo 1</i>	<i>Grupo 2</i>	<i>Grupo 3</i>
<i>-ra</i>	60	85,22	73,5
<i>-se</i>	40	14,78	26,5
N =	60	115	117

**Cuadro 7.** Porcentajes correspondientes a las formas en *-ra* y *-se* por grupos de edad en PRESEGAL. Elaboración propia.

Destacan aquí dos factores distintos. En primer lugar, la frecuencia de las formas simples en *-ra* es mucho más reducida en el grupo de edad más joven, mientras que los otros dos grupos de edad están a una distancia

porcentual considerable, especialmente el intermedio. En segundo término, el grupo más joven es también el que muestra una frecuencia de uso de ambas formas más bajo: supone únicamente el 20,55% del total registrado; el resto se distribuye de forma bastante equilibrada entre los otros dos (un 39,38% y un 40,07%, respectivamente). Vemos, por tanto, que los usos más altos de las formas simples en *-ra* se dan en las mujeres y en personas de mayor edad, especialmente en el grupo intermedio, mientras que los porcentajes más elevados de formas en *-se* (simples) se dan en el grupo más joven.

La comparación entre los datos obtenidos en las encuestas del nivel universitario de PRESEGAL y los procedentes de otros registros o variedades requiere tener en cuenta un factor adicional que no hemos valorado hasta ahora en los recuentos: la existencia de valores indicativos en las formas en *-ra*, cuya consideración hace que la distribución de frecuencias entre *-ra* y *-se* en los contextos en los que cabe la alternancia presente un peso todavía algo mayor para las formas en *-se*.

En efecto, si retiramos de nuestros recuentos los 53 casos de usos indicativos para tomar en consideración únicamente aquellos en los que podría darse la alternancia de las dos formas, obtenemos las distribuciones de la columna “No ind” del cuadro 8, que suponen una ligera reducción de la distancia entre ambas formas con relación a los datos globales de las columnas “Todos” del mismo cuadro.

	Hombres						Mujeres					
	Edad 1		Edad 2		Edad 3		Edad 1		Edad 2		Edad 3	
	Todos	No ind.	Todos	No ind.	Todos	No ind.	Todos	No ind.	Todos	No ind.	Todos	No ind.
<i>-ra</i>	62,96	60	75,6	75	62,74	57,77	57,57	57,57	90,54	83,33	81,18	77,77
<i>-se</i>	37,03	40	24,39	25	37,25	42,22	42,42	42,42	9,45	16,66	18,18	22,22
N =	27	25	41	40	51	45	33	33	74	42	66	54

**Cuadro 8.** Porcentajes correspondientes a las formas en *-ra* y *-se* (valores totales y valores no indicativos) distribuidos por sexo y edad en los textos de PRESEGAL. Elaboración propia.

En resumen, las formas en *-ra* muestran en los hablantes universitarios de la muestra analizada una frecuencia considerablemente menor a la que se observa en el registro oral de otras variedades del español. Los descensos más acusados se registran entre los hablantes del grupo de edad más joven y entre los hombres de más edad, mientras que las mujeres de edad media presentan la proporción más alta de usos de *-ra*, seguidas de las de mayor edad.

Si no resulta fácil identificar tendencias en la distribución de las formas por grupos de edad y sexo, un análisis más detallado de los datos de que disponemos muestra una realidad aun más heterogénea y, sobre todo, más difícilmente interpretable con las correlaciones sociolingüísticas habituales: es necesario tener en cuenta que contamos únicamente con tres entrevistas para cada una de las casillas, con lo que las peculiaridades individuales pueden producir una impresión falsa en alguna de las agrupaciones. Veamos un par de ejemplos claros que destacan en el cuadro 9<sup>14</sup>. Entre los hombres del grupo más joven (H13), que da un porcentaje conjunto de uso de las formas simples en *-ra* del 62,96%, hay dos informantes que no presentan ningún caso de formas en *-se*. El tercer miembro de este grupo, en cambio, muestra 6 casos de *-ra* y 10 casos de *-se*. El resultado de la agrupación de todos los casos de este grupo da la falsa imagen de que en los hombres jóvenes de la muestra el uso de *-se*, con un 37,03%, supera ampliamente la media del conjunto (24,66%), ya de por sí bastante más alta que la documentada en otras variedades orales del español, cuando lo que sucede en realidad es que hay dos informantes que no utilizan nunca la forma en *-se* y otro que, en cambio, la emplea mucho. Algo parecido sucede en el grupo de las mujeres más jóvenes (M13): dos de las informantes utilizan exclusivamente las formas en *-ra*, mientras que la tercera presenta 14 casos de *-se* frente a solo 1 de *-ra*. Como consecuencia, de nuevo la imagen del subgrupo es la de unas frecuencias relativamente próximas de las dos formas (cf. supra cuadro 8) a pesar de que dos de las tres informantes solo utilizan una de ellas. Es evidente que la agrupación de resultados produce, en casos como el que estamos analizando, una visión deformada de lo que sucede en la realidad: dos tercios de las informantes no emplean nunca la forma *-se* y un tercio la utiliza con carácter casi exclusivo.

<sup>14</sup> Vid. las claves utilizadas para las entrevistas en el apéndice.

	<i>-ra</i>	<i>-se</i>	N =
H13_012	100,00	0,00	8
H13_013	37,50	62,50	16
H13_014	100,00	0,00	3
H23_003	100,00	0,00	20
H23_006	37,50	62,50	16
H23_017	100,00	0,00	5
H33_002	93,75	6,25	16
H33_007	50,00	50,00	14
H33_015	47,62	52,38	21
M13_008	100,00	0,00	10
M13_010	6,67	93,33	15
M13_016	100,00	0,00	8
M23_001	100,00	0,00	18
M23_004	92,16	7,84	51
M23_018	40,00	60,00	5
M33_005	100,00	0,00	17
M33_009	100,00	0,00	9
M33_011	70,00	30,00	40
	75,34	24,66	292

**Cuadro 9.** Porcentajes de las formas simples en *-ra* y *-se* en las entrevistas del nivel sociocultural 3 de PRESEGAL. Fuente: PRESEGAL. Elaboración propia.

Estas fuertes discrepancias entre informantes procedentes del mismo subgrupo (sexo y edad) se extienden a lo largo y ancho de todas las encuestas. Frente a la impresión, basada en datos cuantitativos, de porcentajes elevados para las formas *-se*, el análisis individualizado muestra que 9 de las 18 encuestas no presentan ningún caso de *-se* y otras dos muestran porcentajes de *-ra* superiores al 90%. De las siete restantes, tres tienen porcentajes de *-se* superiores al 60% y las demás se sitúan en una zona intermedia. Para paliar el efecto que las entrevistas con gran número de usos de estas formas

producen sobre los resultados generales, podemos optar por una visión diferente, la basada en las medias de porcentajes de uso. Evidentemente, esa aproximación presenta el grave inconveniente de que atribuye el mismo peso a entrevistas que tienen números de ejemplos muy discrepantes entre sí, pero quizá en este caso sea un precio preferible al que supone la otra aproximación. De todos modos, a pesar de la evidente diferencia de enfoques, la media de porcentajes correspondientes a las formas simples en *-ra* es 74,03, muy próximo al 73,34% que hemos obtenido con los datos generales, aunque ahora podemos matizarlo con la indicación de que la desviación estándar de los porcentajes individuales de las formas en *-ra* se eleva a 27,23. Veamos las cifras que resultan de los cruces desde esta segunda perspectiva, que son las que aparecen en los cuadros 10 y 11.

	<i>Hombres</i>	<i>Mujeres</i>
<i>-ra</i>	74,04	78,76
<i>-se</i>	25,96	21,24
N =	119	173

**Cuadro 10.** Medias de porcentajes correspondientes a las formas en *-ra* y *-se* por sexos en PRESEGAL. Elaboración propia.

	<i>Grupo 1</i>	<i>Grupo 2</i>	<i>Grupo 3</i>
<i>-ra</i>	74,03	78,28	76,89
<i>-se</i>	25,97	21,72	23,11
N =	60	115	117

**Cuadro 11.** Medias de porcentajes correspondientes a las formas en *-ra* y *-se* por grupos de edad en PRESEGAL. Elaboración propia.

La comparación con lo que figura en los cuadros 6 y 7 muestra que el efecto más marcado del cambio de perspectiva consiste en la notable elevación que se observa en los porcentajes de uso de las formas simples en *-ra* en las casillas que estaban más bajas: hombres por un lado y grupo más joven por otro, con la consiguiente aproximación de los resultados que figuran en las diferentes casillas, pero sigue siendo necesario tener en cuenta que existen fuertes diferencias individuales en el uso de las dos formas, diferencias que cualquier medida de tendencia central oculta en un grado difícil de tolerar debido tanto a la magnitud de las divergencias como al escaso tamaño de la muestra utilizada.

### 3.3. Usos indicativos de las formas en *-ra*

Como ya hemos indicado, el español de Galicia posee usos plenamente indicativos de las formas en *-ra*, equivalentes a los del español general *había amado* –*vieran* en (3), *entendiera* en (4) –, *amé* o *amaba* –*llorara* y *dijeran* en (3), y *fuera* en (5) – y *amaría* –como *debiera* en (6).

- (3) mi abuelo [...] bueno lloraba mucho *llorara* mucho / cuando le *dijeran* que *vieran* a mi tía / o sea la hermana de mi madre fumando // M23\_001
- (4) pero después cuando tuve que dar clase // eeh como ya habían pasado años / y ya se ya había producido o sea ya se el proceso de maduración que de que te dan años de experiencia // pues / resulta que que que me encontré / estudiando cosas // que nunca *entendiera* / o sea yo no // en la carrera no entendía / eeh cosas que aprendí pero no las entendí ¿eh? H33\_007
- (5) porque eran unos unos negreros hombre / el antiguo mm director *fuera* un negrero / y te hacían cada cada una que no sé / M33\_011
- (6) yo no sé si si no equivoqué la carrera / yo creo que *debiera* haber est estudiado Historia / H33\_007

Solo el último de ellos, reducido a verbos modales, puede encontrarse con cierta extensión en otras variedades del español. En las 18 entrevistas analizadas, hemos encontrado 53 casos de alguno de estos tres usos, lo cual supone el 24,09% del total de formas simples en *-ra*. Como hemos venido haciendo hasta ahora, en los cuadros 12 y 13 incluimos la distribución porcentual de estos empleos cruzados con sexo y grupo de edad.

	<i>Hombres</i>	<i>Mujeres</i>	<i>N =</i>
<i>-ra</i>	16,98	83,02	53

**Cuadro 12.** Porcentajes correspondientes a usos indicativos de las formas en *-ra* por sexos en PRESEGAL. Elaboración propia.

	<i>Grupo 1</i>	<i>Grupo 2</i>	<i>Grupo 3</i>	<i>N =</i>
<i>-ra</i>	3,77	62,26	33,96	53

**Cuadro 13.** Porcentajes correspondientes a usos indicativos de las formas en *-ra* por grupos de edad en PRESEGAL. Elaboración propia.



Teniendo en cuenta lo que ya hemos visto sobre los datos generales, no resultará inesperada la comprobación de que estos usos parecen estar concentrados en las mujeres y apenas tienen presencia en el grupo más joven. Y tampoco será sorpresa saber que la agrupación de datos disfraza en buena parte lo que sucede en la realidad. No hay ningún ejemplo de estos usos en 3 informantes masculinos y 4 femeninos, lo cual significa que está totalmente ausente en el 38,88% de las entrevistas. Además, en 5 hombres y 2 mujeres encontramos un único caso. Por tanto, estos empleos no se dan o aparecen solo una vez en 14 de las 18 entrevistas. Como muestra el cuadro 14, el peso de estos usos se concentra en 3 informantes femeninas, que aportan el 79% de los casos registrados, y 25 de ellos (el 47,17% del total) se debe a una integrante del grupo de edad intermedio.

	Usos indicativos de -ra
H13_012	1
H13_013	1
H13_014	0
H23_003	1
H23_006	0
H23_017	0
H33_002	1
H33_007	4
H33_015	1
M13_008	0
M13_010	0
M13_016	0
M23_001	7
M23_004	25
M23_018	0
M33_005	1
M33_009	1
M33_011	10
Total	53

**Cuadro 14.** Casos de formas simples en -ra con valor indicativo en el nivel sociocultural 3 de PRESEGAL. Elaboración propia.

Dado que la ausencia de formas en *-ra* con valor indicativo podría proceder también de la escasez relativa de las condiciones sintácticas y discursivas adecuadas para su aparición, resultará útil contrastar, para cada informante, el número de casos de *-ra* con valor de situación anterior a un punto anterior al origen (esto es, de antepretérito o antecopretérito) con los de *había + participio* con este mismo valor temporal. Es lo que muestra el cuadro 15:

	<i>-ra</i>	<i>había + part.</i>	<i>N =</i>
H13_012	0,00	100,00	1
H13_013	33,33	66,67	3
H13_014	0,00	100,00	12
H23_003	0,00	100,00	3
H23_006	0,00	100,00	11
H23_017	0,00	100,00	3
H33_002	9,09	90,91	11
H33_007	50,00	50,00	4
H33_015	7,69	92,31	13
M13_008	0,00	100,00	3
M13_010	0,00	100,00	17
M13_016	0,00	100,00	6
M23_001	14,29	85,71	14
M23_004	33,33	66,67	18
M23_018	0,00	0,00	0
M33_005	10,00	90,00	10
M33_009	5,56	94,44	18
M33_011	54,55	45,45	11
<i>N =</i>	21	137	158

**Cuadro 15.** Porcentajes de formas en *-ra* con valor de antecopretérito y formas del tipo *había + participio* con el mismo valor temporal en las 18 entrevistas de PRESEGAL. Elaboración propia.

Las 14 entrevistas en las que aparece un máximo de un caso con este valor (el de antecopretérito) son ahora únicamente 2, de modo que no se trata de la inexistencia de esa relación temporal en las situaciones descritas. El uso de las formas en *-ra*, que consideramos característico del español de Galicia (y del gallego), se da únicamente en el 13,9% del total de casos con este valor. En efecto, hay 8 entrevistas que utilizan exclusivamente la forma compuesta. De las restantes, 8 muestran mayor peso de la forma compuesta y solo 2 (H33\_007 y M33\_011) emplean las formas en *-ra* en aproximadamente la mitad de los casos.

En este caso, pues, parece que la forma propia del español general se impone sobre la característica del español de Galicia en los informantes que constituyen nuestra muestra. Para tomar en consideración todas las caras del fenómeno, hemos analizado también los casos en los que las formas en *-ra* tienen valor de pretérito o copretérito y, a su lado, los de *había + participio* con este mismo valor, que, como hemos señalado, es un fenómeno inducido por la idea de que las formas en *-ra* son rechazables y deben ser sustituidas por *había + participio*. El fragmento de (7) ofrece estos usos de pretérito tanto de la forma compuesta (*había ido*) como de la forma en *-ra* (*fuera, encantara*). En (8) tenemos un caso con valor de copretérito (*había vivido*).

- (7) [está hablando de actividades] voy a pocas no voy no suelo / quitarle partido seguramente a lo que tenemos ahí / porque sí hacen cosas ¿eh? / sí hacen cosas interesantes / hm yo *había ido* en el colegio de / en el colegio de los niños cuando iban / a La Salle / pues *fuera* a uno de flo de técnica de flores secas porque las manualidades me gustan mucho también / y de cua de hecho cuad esos cuadros por aquí están por ahí los hice yo / est estes que están ahí de flores secas ¿sabes? / y me gusta mucho / esas eeh y y y se me dan bien / la verdad / y *fuera* al y me *encantara* el curso de flores secas M23\_004
- (8) entonces eeh / mi marido siempre le contaba que claro que cuando *había vivido* en Andalucía en el pueblo // pues tenían tirachinas / entonces bueno allá / un día // eeh / le hizo uno a mi hijo pequeño // y le dijo / no lo lledes al colegio / esto no es para llevar al colegio / M23\_001

Así pues, la forma *había + participio* ocupa no solo el terreno que le corresponde en español general, sino también el que corresponde a las formas en *-ra* en gallego y en español de Galicia. Los datos figuran en el cuadro 16.

	<i>-ra</i>	<i>había + part.</i>	N =
H13_012	0,00	0,00	0
H13_013	0,00	100,00	1
H13_014	0,00	0,00	0
H23_003	100,00	0,00	1
H23_006	0,00	100,00	1
H23_017	0,00	100,00	2
H33_002	0,00	100,00	5
H33_007	50,00	50,00	2
H33_015	0,00	100,00	1
M13_008	0,00	0,00	0
M13_010	0,00	100,00	4
M13_016	0,00	100,00	2
M23_001	83,33	16,67	6
M23_004	51,35	48,65	37
M23_018	0,00	0,00	0
M33_005	0,00	100,00	4
M33_009	0,00	100,00	2
M33_011	100,00	0,00	4
N =	30	42	72

**Cuadro 16.** Porcentajes de formas simples en *-ra* con valor de pretérito o copretérito y formas del tipo *había + participio* con el mismo valor temporal en las 18 entrevistas de PRESEGAL. Elaboración propia.

De los datos que figuran en el cuadro 16 hay que destacar, en primer lugar, que los usos de alguna de las dos formas estudiadas con valor de pretérito o copretérito se dan en 14 de las 18 entrevistas, lo cual supone un

porcentaje importante dada la peculiaridad del fenómeno analizado. En el paso siguiente, sorprende que la forma *había + participio*, cuyo uso con este valor se explica mediante el mecanismo de equiparación con las formas en *-ra* y el deseo de evitar el uso de estas, lo cual la hace penetrar en un territorio nuevo, tenga aquí el 58,33% de los empleos con estos valores. Como es de esperar, los usos individuales se mueven a lo largo de todo el espectro de posibilidades, desde la M33\_011, con el 100% de los 4 casos correspondientes a formas en *-ra*, al igual que H23\_003, pero este con un único caso, hasta, al otro extremo, los nueve informantes en los que el 100% corresponde a la forma compuesta, aunque tres de ellos presentan un único caso. En medio, la siempre destacada M23\_004, que presenta 37 casos con este valor en alguna de las dos formas (el 51,39% del total registrado) y los distribuye entre ellas de forma casi igualitaria.

#### 4. PERSPECTIVA GENERAL Y CONCLUSIONES

Evidentemente, la variedad que llamamos habitualmente «español de Galicia» surge como consecuencia del contacto entre gallego y español en un conjunto de circunstancias sociales relativamente bien conocidas<sup>15</sup>, que dan como resultado el aprendizaje del español a partir del gallego, con lo que se produce la convivencia de ambas lenguas en los mismos hablantes, muchas veces con marcada diferenciación de ámbitos de uso. Es decir, se pasa, no en general, sino en los individuos, de una situación de diglosia de adscripción a otra de diglosia funcional (cf. Rojo 1985) y también, con cierta frecuencia hasta hace unos cuantos años, al abandono de la primera. La proximidad de ambas lenguas crea un terreno bien abonado para las interferencias, pero hay que tener en cuenta también que muchos hablantes de español de Galicia no son hablantes originarios de gallego que hayan aprendido español, sino personas que han tenido como lengua primera una variedad del español que contenía ya interferencias resultantes de aprendizajes realizados una o dos generaciones antes. En estos casos no se trata de interferencias, sino de casos de integración (cf. Rojo 2004: 1090 y sigs.). Así se entiende bien la sensación que produce en esas personas el darse cuenta de que su español difiere de, por ejemplo, el propio del centro peninsular en algunos aspectos bastante marcados como, por no citar más

<sup>15</sup> Vid. la bibliografía mencionada en la nota 4 de Rojo 2004.

que fenómenos gramaticales, la baja frecuencia de formas de antepresente, antefuturo o antepostpretérito, el uso de la perífrasis *dar + participio*, la extensión de la construcción *tener + participio* y muchos otros, entre los que destaca el empleo de las formas simples en *-ra* con valores que no presenta en otras variedades del español.

En la Galicia tradicional, el proceso de ingreso al español (de Galicia) se producía en paralelo a un proceso general de ascenso en la escala social y el aumento de la presencia del español (en diferentes variedades) no solo en el sistema educativo, la esfera religiosa, los medios de comunicación y los ámbitos oficiales, sino también en los familiares. De ahí que, entre los hablantes de español de Galicia en la actualidad, podamos encontrar desde quienes han llegado al español desde el gallego en un proceso personal de aprendizaje no reglado (y tienen, por tanto, interferencias) hasta quienes lo han tenido ya como primera lengua, adquirida en el seno de una familia que abandonó el uso del gallego una o dos generaciones antes. A su lado, por supuesto, hablantes con español de otras zonas en su familia, personas que han nacido o vivido fuera, etc. Tener todo esto en cuenta y comprender la consideración de todas estas posibilidades generales permite entender el hecho de que los fenómenos que hemos estudiado aquí no muestren correlaciones claras con las variables sociales consideradas habitualmente. En realidad, no estamos ante un proceso de cambio lingüístico, ni siquiera en la fase que corresponde a los indicadores de Labov, pero eso no significa que estas características especialmente marcadas no puedan ser consideradas estereotipos. Los que intervienen son fenómenos en buena parte estrictamente personales, aunque, por supuesto, la estratificación social pesa (o ha pesado) sobre las circunstancias desencadenantes del proceso.

Las formas en *-ra*, que constituyen en sí mismas un fenómeno del mayor interés en todos los ejes de variación, presentan en el español de Galicia algunos aspectos adicionales que hemos tratado de analizar a partir de los datos proporcionados por las 18 encuestas del nivel sociocultural 3 de PRESEGAL en Santiago de Compostela. En la consideración general, las frecuencias relativas de las formas en *-ra* y *-se* presentan en la muestra estudiada un predominio de *-ra* considerablemente menor al observado en la mayor parte de las variedades, tanto en general como en el registro oral, a pesar de que el proceso de penetración de las formas en *-ra* en valores subjuntivos se da también en gallego y, por tanto, no encuentra resistencia sistémica para darse en el español de Galicia. Con los datos de que disponemos en este momento, la mayor o menor frecuencia relativa de las formas en *-ra*

no parece estar correlacionada con el sexo o la edad y encontramos tanto 9 informantes que utilizan únicamente las formas en *-ra* como 4 informantes que presentan formas en *-se* en porcentajes superiores al 60% del total (cf. cuadro 9), muchos de ellos con un número importante de casos que impiden pensar en factores exclusivamente casuales.

Encontramos, pues, dos tendencias bien marcadas en los extremos. De una parte, la más extendida en el conjunto de las variedades del español, especialmente en registros orales, consistente en ampliar el uso de las formas en *-ra* hasta provocar la desaparición de las formas en *-se*, fenómeno que se ha producido en la mitad de las entrevistas de la muestra. En el otro extremo, la que choca con la más habitual en español, pero encaja mejor con lo que sucede en gallego actual, consistente en que el deslizamiento de las formas en *-ra* hacia valores subjuntivos no esté bloqueado, pero sí reducido en frecuencia, de modo que las formas en *-se* conservan un porcentaje de uso amplio, incluso mayoritario, en los contextos en los que es posible la alternancia. La hipótesis es tentadora: la utilización de *-ra* con valores indicativos debería favorecer la conservación del uso de *-se* con valores subjuntivos. Y, en efecto, vemos que el porcentaje de *-se* sube, tanto con respecto a datos generales del español oral como con relación a la media de PRESEGAL, en algunos hablantes que presentan usos indicativos de *-ra* (H13\_013, H33\_007, H33\_015, M33\_011), hasta igualar o superar incluso la frecuencia de *-ra* en tres de estos cuatro hablantes. Sin embargo, el empleo de *-se* no deja de ser, en términos de frecuencia relativa global, la opción marcada frente a *-ra*, de modo que hay otros 5 hablantes que usan *-ra* con valor indicativo (H13\_012, H23\_003, M23\_001, M33\_005, M33\_009) y no producen ningún caso de *-se*, más otros 2 que presentan porcentajes de *-se* bastante inferiores a la media (H33\_002, 6,25% de *-se*; M23\_004, 7,84%).

Las dos opciones apuntadas parecen darse, pues, con independencia de que existan o no usos indicativos de *-ra*. Los 5 hablantes que presentan formas en *-ra* con valores indicativos y no tienen formas en *-se* (H13\_012, H23\_003, M23\_001, M33\_005, M33\_009) muestran la posibilidad de llevar la extensión de uso de las formas en *-ra* hasta sus últimas consecuencias: conservan plenamente los valores indicativos y se extienden en los subjuntivos hasta producir la desaparición de *-se* en estas muestras. Es esta una posibilidad novedosa, puesto que difiere de la que encontramos en otras variedades del español en el mantenimiento pleno de los usos indicativos y, al tiempo, de la que se da en gallego actual en la ausencia de las formas en *-se*.

La cuestión complementaria es igualmente importante. ¿Por qué no se da esta línea en todos los casos y, en cambio, encontramos hablantes que presentan frecuencias inesperadamente altas de formas en *-se*, combinadas tanto con la utilización de formas en *-ra* con valor indicativo como con la ausencia total de estos usos? La primera posibilidad -usos indicativos de *-ra* y predominio de *-se* en los empleos subjuntivos- es similar a la que se da en gallego. La segunda -ausencia de usos indicativos para las formas en *-ra* y predominio de *-se* en los valores subjuntivos- choca tanto con la general en gallego por la ausencia de los valores indicativos como con la propia del español, que ha perdido la mayor parte de los usos indicativos de las formas en *-ra*, pero ha desplazado esta forma hacia usos subjuntivos que reducen fuertemente o incluso anulan totalmente la utilización de las formas en *-se*. A esta segunda opción pueden contribuir al menos dos factores distintos, que actúan conjuntamente en la misma dirección. Desde el ámbito propio del español de Galicia, puede pensarse que la influencia de la estigmatización que se da sobre los usos indicativos de las formas en *-ra* se extienda a todos los usos de estas formas. En paralelo, la influencia de la distribución habitual en gallego, que utiliza mayoritariamente formas en *-se* para usos subjuntivos, y también del rasgo añadido de mayor elegancia que suele atribuírsele en español, pueden contribuir conjuntamente a que las formas en *-se* tengan mayor frecuencia de uso o incluso sean exclusivas en los empleos subjuntivos. Profundizar en las causas de esta opción requiere de una parte aumentar la muestra tanto en extensión como en profundidad y también explorar lo que sucede con los diferentes contextos de uso en los que es posible la alternancia (condicionales, casos de rección temporal, subjuntivo regido, etc.), factores todos ellos que caen fuera del ámbito de este trabajo.

El último punto de este apasionante proceso implica, como hemos visto, a las formas del tipo *había + participio*, que presenta en algunos hablantes valores de pretérito o copretérito, ejemplificados antes en (7) y (8), y también en los siguientes fragmentos:

- (9) yo no he estado en Egipto o sea tengo una hija que sí  
estuvo la verdad le *había encantado* / H33\_002
- (10) mmm bueno en / sí en octavo de EGB *habíamos ido* a  
/ a Navarra a esquiar con los / con los compañeros de  
clase y los profesores / y la verdad que de ese / de esa  
excursión tengo muy / muy buen recuerdo hm / y luego  
en el instituto hubo otras pero no / no pude ir H13\_013



Como hemos visto en el cuadro 16, son valores relativamente frecuentes: 30 sobre 53 casos de formas en *-ra* tienen valor de pretérito o copretérito, a los que hay que añadir los 42 con este mismo valor de formas del tipo *había* + *participio* (que son nada menos que el 23,46% de los usos de esta forma). El origen del valor de pretérito de *había*+*participio* reside en la asunción por parte de sus usuarios de una equivalencia estricta entre los usos indicativos de *-ra* y los de la forma compuesta. La asociación de tales usos de *-ra* con el sistema verbal del gallego y la conciencia de que no son propios de *-ra* en el estándar español conducen a la elección de la forma compuesta para evitar lo que en términos normativos se considera un error en español por interferencia del gallego. Se transfieren entonces a *había*+*participio* todos los valores indicativos de *-ra* y no solo el de antecopretérito, de tal manera que los usos de pretérito de *había*+*participio* constituyen un efecto indirecto, y por tanto menos ostensible, de la constelación de valores temporales expresados por *-ra* en español de Galicia (y por *-ra* en gallego).

Resulta, pues, que la manifestación de los valores indicativos de *-ra* en la muestra no se reduce a los 53 casos contabilizados y a los hablantes correspondientes (que son 11 de 18), sino que se extiende a los usos de pretérito de *había*+*participio* y, consiguientemente, a los hablantes que los producen. Algunos de ellos coinciden con los 11 usuarios de *-ra* indicativo, pero hay también otros “nuevos”: H23\_006, H23\_017, M13\_010 y M13\_016. De esta manera, solo 3 informantes de los 18 presentan un sistema ajeno, aparentemente, a la influencia de *-ra* indicativo.

Al constatar la relación de los valores temporales de *-ra* y de *había*+*participio* en la línea que acabamos de exponer, cobran sentido también los datos de uso de la forma *-se*. Veíamos que algunos hablantes sin usos indicativos de *-ra* presentaban elevados porcentajes de *-se*, lo cual parecía echar por tierra la hipótesis de que el aumento de esta segunda forma, de muy baja frecuencia en español oral general, estuviese condicionada por una especialización de funciones en contraste con *-ra*. Sin embargo ahora podemos confirmar que la alta proporción de *-se* es en todos los casos la consecuencia de una misma tendencia diferenciadora frente a los usos indicativos de *-ra*. Todos los hablantes que superan la media de usos de *-se* en el corpus muestran de forma directa —a través de los usos indicativos de *-ra*— o de manera indirecta —a través de los usos de pretérito de *había*+*do*— los efectos de compartir un mismo sistema básico de relaciones modo-temporales, sistema que muestra diferencias y concomitancias tanto con otras variedades del español, por una parte, como con el sistema común del gallego, por otra.

## REFERENCIAS A CORPUS Y OTROS RECURSOS ELECTRÓNICOS MENCIONADOS EN EL TEXTO

- BDS: *Base de datos sintácticos del español actual*.  
<<http://www.bds.usc.es/>>.
- CE: *Corpus del Español: 100 million words, 1200s-1900s*.  
<<http://www.corpusdelespanol.org>>.
- CORDE: Real Academia Española: *Corpus diacrónico del español*.  
<<http://www.rae.es/recursos/banco-de-datos/corde>>.
- CORGA (anotado): Centro Ramón Piñeiro para a investigación en Humanidades: *Corpus de referencia do galego actual (etiquetado)*  
<<http://corpus.cirp.es/corgaetq/>>.
- CREA: Real Academia Española: *Corpus de referencia del español actual*.  
<<http://www.rae.es/recursos/banco-de-datos/crea>>.
- FREELING: *An Open Source Suite of Language Analyzers*  
<<http://nlp.lsi.upc.edu/freeling/>>.
- PRESEEA: *Proyecto para el estudio sociolingüístico del español de España y de América* <<http://preseea.linguas.net/>>.
- PRESEGAL: *Proyecto para el estudio sociolingüístico del español de Galicia*. <<http://gramatica.usc.es/proxectos/presegal/>>.

### DATOS DE LAS ENCUESTAS DE PRESEGAL UTILIZADAS

	20-34 años	35-54 años	>54 años
Mujeres	SCOM_M13_008	SCOM_M23_001	SCOM_M33_005
	SCOM_M13_010	SCOM_M23_004	SCOM_M33_009
	SCOM_M13_016	SCOM_M23_018	SCOM_M33_011
Hombres	SCOM_H13_012	SCOM_H23_003	SCOM_H33_002
	SCOM_H13_013	SCOM_H23_006	SCOM_H33_007
	SCOM_H13_014	SCOM_H23_017	SCOM_H33_015



## REFERENCES

- Álvarez, Rosario, Henrique Monteagudo y Xosé Luís Regueira (1995): *Gramática galega*. Vigo: Galaxia, 6<sup>a</sup> edición.
- y Xosé Xove (2002): *Gramática da lingua galega*. Vigo: Galaxia.
- Bassols de Climent, Mariano (1956): *Sintaxis latina*. Madrid: CSIC.
- Bechara, Evamildo (1999): *Moderna Gramática portuguesa*. Rio de Janeiro: Lucerna, 37<sup>a</sup> ed. revista e ampliada.
- Becker, Martin G. (2008): “From temporal to modal. Divergent fates of the temporal latin synthetic pluperfect in Spanish and Portuguese”, en Ulrich Detges y Richard Waltereit (eds.): *The Paradox of Grammatical Change. Perspectives from Romance*. Amsterdam: John Benjamins, pp. 147-179.
- Bejarano, Virgilio (1962): “Sobre las dos formas del imperfecto de subjuntivo y el empleo de la forma en -se con valor de indicativo. *Strenae. Estudios de filología e historia dedicados al profesor Manuel García Blanco* (=Acta Salmanticensia, 16), Salamanca, pp. 77-86.
- Blas Arroyo, José Luis (1991): “Problemas teóricos en el estudio de la interferencia lingüística”, *Revista Española de Lingüística* 21 (2), pp. 265-289.
- (1993): *La interferencia lingüística en Valencia*. Castellón: Universitat Jaume I.
- Cunha, Celso e Lindley Cintra (1984): *Nova gramática do português contemporâneo*. Lisboa: João Sa da Costa, 1994<sup>10</sup>.
- De Sterck, Goedele (2000): *Gramática: registros y áreas geográficas en lingüística: valores y usos de las formas verbales en -ra, -se, -ría y -re*. Vol. II, 7 de la serie editada por Josse De Kock *Gramática española: enseñanza e investigación*. Salamanca: Universidad de Salamanca.

- Freixeiro Mato, Xosé Ramón (2000): *Gramática da lingua galega. II Morfosintaxe*. Vigo: A Nosa Terra.
- Girón Alconchel, José Luis (2000): “Análisis del discurso y cambio lingüístico (sobre la historia de *cantara* indicativo)”, en José Jesús Bustos Tovar, Patrick Charaudeau, José Luis Girón Alconchel, Silvia Iglesias Recuero y Covadonga López Alonso (eds.): *Lengua, discurso, texto. (I Simposio internacional de análisis del discurso)*. Madrid: Visor, pp. 309-322.
- González Ollé, Fernando (2012): *Continuidad histórica ininterrumpida de la forma -ra de indicativo. Tradiciones discursivas y sintaxis*. Navarra: Eunsa.
- Hermerén, Ingrid (1992): *El uso de la forma en ra con valor no-subjuntivo en el español moderno*. Lund University Press.
- NGLE: Real Academia Española / Asociación de Académicas de la lengua española: *Nueva gramática de la lengua española*. Madrid: Espasa, vol I.
- Nowikow, Wiaczeslaw (1991): “Según dijera, después que llegara... ¿un americanismo?”, en César Hernández Alonso et al. (eds.): *El español de América. Actas del III Congreso Internacional del español de América*. Valladolid: Junta de Castilla y León, vol. I, pp. 537-540.
- Padró, Lluís y Evgeny Stanilovsky (2012): “FreeLing 3.0: Towards Wider Multilinguality”, en *Proceedings of the Language Resources and Evaluation Conference (LREC 2012) ELRA*. Estambul, mayo 2012.
- Pollán, Celia (2001): “The expression of pragmatic values by means of verbal morphology: A variationist study”, *Language Variation and Change* 13, pp. 59-89.  
<http://dx.doi.org/10.1017/S0954394501131030>
- (2003): “The morphological expression of pragmatic values in oral and written Galician”, en Manuel Fernández Ferreiro y Fernando Ramallo (eds.): *Sociolinguistics in Galicia: Views on diversity, a diversity of views [= Estudios de Sociolingüística, 3/2 (2002) y 4/1 (2003)]*, pp. 113-138.
- RAG / ILG (2003): Real Academia Galega / Instituto da Lingua Galega: *Normas ortográficas e morfolóxicas do idioma galego*. Santiago de Compostela, 18<sup>a</sup> ed.
- Royo, Guillermo (1974): «La temporalidad verbal en español», *Verba* 1, pp. 68-149.

- (1985): «Diglosia y tipos de diglosia», en *Philologica Hispaniensi*. *In honorem M. Alvar*, Madrid: Gredos, vol. II, pp. 603-617.
  - (1996): “Sobre la distribución de las formas *llegara* y *llegase* en español actual”, en Manuel Casado Velarde *et al.* (eds.): *Scripta Philologica in memoriam Manuel Taboada Cid*. A Coruña: Universidade da Coruña, II, pp. 677-691.
  - (2004): «El español de Galicia», en Rafael Cano (coord.): *Historia de la lengua española*. Barcelona: Ariel, pp. 1087-1101, 2005<sup>2</sup>.
  - (2010): “Sobre codificación y explotación de corpus textuales: Otra comparación del *Corpus del español* con el CORDE y el CREA”. *Lingüística* 24, pp. 11-50.
  - (2011a): “De nuevo sobre la frecuencia de las formas *llegara* y *llegase*”, en Jörn Albrecht y Frank Harslem (eds.): *Heidelberger Spätlese. Ausgewählte Tropfen aus verschiedenen Lagen der spanischen Sprach- und Übersetzungswissenschaft. Festschrift anlässlich des 70. Geburtstages von Prof. Dr. Nelson Cartagena*. Bonn: Romanistischer Verlag, pp. 161-182.
  - (2011b): “Me pidieron que reseñara~reseñase el libro que ?publicara / \*publicase Bosque en 1980”, en M. Victoria Escandell Vidal, Manuel Leonetti y Cristina Sánchez López (eds.): *60 problemas de gramática dedicados a Ignacio Bosque*. Akal: Madrid, pp. 213-219.
  - y Emilio Montero Cartelle (1983): *La evolución de los esquemas condicionales. (Potenciales e irreales desde el Poema del Cid hasta 1400)*, anejo 22 de Verba. Santiago de Compostela: Universidade de Santiago de Compostela.
  - y Alexandre Veiga (1999): “El tiempo verbal. Las formas simples”, cap. 44 de Ignacio Bosque y Violeta Demonte (eds.): *Gramática descriptiva de la lengua española*, Madrid: Espasa-Calpe, pp. 2867-2934.
- Santalla del Río, María Paula y Susana Sotelo Docío (2006): “Constitución y etiquetación del corpus CESAR”, en Juan de Dios Luque (ed.): *Actas del V Congreso andaluz de Lingüística general. Homenaje al profesor José Andrés de Molina Redondo*. Granada: Universidad de Granada, I, pp. 369-382.
- Solà, Joan, Maria-Rosa Lloret, Joan Mascaró, Manuel Pérez Saldanya (dirs.) (2002): *Gramàtica del català contemporani*. Barcelona: Empúries, 3 vols.

- Veiga, Alexandre (1996): *La forma verbal española cantara en su diacronía*. Santiago de Compostela: Universidade de Santiago de Compostela.
- (2005): “Falsos castellanimos en gallego. Anotaciones a la luz de la lingüística diacrónica”. *Res Diachronicae Virtual* 4, pp. 23-50.
  - (2006): “Las formas verbales subjuntivas. Su reorganización modo-temporal”, en Concepción Company Company (dir.): *Sintaxis histórica de la lengua española*. México, D.F.: FCE, I, pp. 93-240.
  - (2011): “Algo más sobre la doble reorganización temporal en la diacronía del sistema verbal español”, en Carsten Sinner, José Luis Ramírez Luengo y María Jesús Torréns Álvarez (coords.): *Tiempo, espacio y relaciones espacio-temporales dentro de la perspectiva de la lingüística histórica*. San Millán de la Cogolla: Instituto de historia de la lengua, pp. 297-330.
- Xove, Xosé (1986): “As formas en *-ra* e *-se* na obra poética galega de Rosalía”, en *Actas do Congreso internacional de estudos sobre Rosalía de Castro e o seu tempo*. Santiago de Compostela: Universidade de Santiago de Compostela, vol. III, pp. 133-149.



## **Efectos del contacto de lenguas en el castellano de Mallorca: una perspectiva histórica\***

Andrés Enrique-Arias  
*Universitat de les Illes Balears*

**Abstract:** This study uses two phenomena of the Spanish in contact with Catalan in Mallorca (the verb *pedir* ‘ask for something’ with the semantic and syntactic structure of *preguntar* ‘ask a question’ and the preponderance of synthetic *cantaré* over periphrastic future *voy a cantar*) as case studies in order to emphasize the relevance of a historical perspective in the study of contact varieties. The analysis reveals that (i) these features are attested in Mallorcan Spanish at least since the 1700s and (ii) they had some sort of existence in monolingual varieties of Spanish at the time when the language was introduced to the island. The historical data suggest that these features of Mallorcan Spanish are not, as traditionally believed, contact-driven innovations, but the result of the preservation of structures that are recessive in monolingual Spanish but in Mallorca have been reinforced by the existence of a parallel Catalan structure.

**Keywords:** Language contact, language variation and change, historical linguistics, morphosyntax, corpus linguistics.

---

\* Investigación financiada con una beca del Programa Puente de La Caixa (2014).





## 1. INTRODUCCIÓN

En los estudios sobre lenguas en contacto llevados a cabo en diferentes áreas del mundo hispánico se plantean con frecuencia cuestiones en torno al conocido debate sobre el alcance del cambio lingüístico inducido por la transferencia: si cualquier elemento puede ser transferido de una lengua a otra; en qué medida y de qué manera son los factores externos (la situación sociolingüística) o internos (la estructura y la tipología de las lenguas mismas) responsables de la dirección e intensidad de la transferencia; o cuáles son los efectos concretos que produce el contacto lingüístico. En lo que respecta a esta última cuestión, una de las afirmaciones más comunes que aparecen en los estudios de lingüística diacrónica es que el contacto de lenguas tiene como consecuencia prácticamente inevitable el cambio lingüístico. La razón es que el bilingüismo propicia toda una serie de fenómenos que dan lugar a cambios de diversa especie: intercambio de códigos, préstamos, simplificación de categorías gramaticales y léxicas, desarrollo de construcciones perifrásticas, regularización de patrones lingüísticos, entre otros. Estos cambios se intensifican en situaciones en que hay un bilingüismo de sociedad desequilibrado en favor de una de las lenguas en contacto: cuando hay una lengua que es socialmente dominante y una lengua subordinada, la situación sociolingüística que caracteriza a esta última, como la falta de presión normativa o la restricción en los contextos de uso de la lengua, contribuyen a intensificar y extender las variantes innovadoras. El resultado sería la aceleración de tendencias de cambio que se encuentran en un nivel de desarrollo menos avanzado en variedades monolingües de la misma lengua (véase por ejemplo, Silva-Corvalán 1994: 7, 119, 130).

Al mismo tiempo son relativamente pocas las investigaciones que consideran la situación opuesta, es decir, la posibilidad de que la convivencia de dos lenguas en la misma comunidad de habla haga que en una de ellas se conserven variantes que están retrocediendo en variedades monolingües de esa misma lengua. Esta es precisamente la situación descrita en varios estudios sobre la evolución histórica del castellano hablado en zonas de habla catalana, donde debido a la influencia del catalán se conservan con notable vitalidad algunos rasgos lingüísticos que están retrocediendo o incluso han desaparecido en el español monolingüe (véase el estudio de Blas Arroyo 2007 con hablantes de Castellón, o el trabajo de Enrique-Arias 2010 sobre el castellano en contacto con el catalán en Mallorca).

La sociolingüística variacionista y los estudios experimentales de bilingüismo y adquisición de segundas lenguas han estudiado de manera fructífera las situaciones de cambio inducido por el contacto —en particular la transmisión y adquisición de variantes innovadoras— mediante la observación de diferentes grupos generacionales sincrónicamente. Aun dando por supuesto el valor de estos planteamientos metodológicos no cabe duda de que a la hora de establecer que en una variedad dada se ha producido un cambio inducido por el contacto es necesario examinar la historia de las lenguas implicadas e investigar si los rasgos propuestos como resultado de la interferencia estaban presentes en la variedad en cuestión *antes* del contacto, algo sobre lo que Thomason (2001: 93-94) llamaba la atención en los siguientes términos:

[W]e must prove that the shared features – the proposed interference features – were not present in the receiving language before it came into close contact with the source language. That is, we have to prove that the receiving language has changed by innovating these features.

Por ello, la incorporación de una perspectiva histórica es un componente relevante y enriquecedor en los estudios de contacto de lenguas: solo de esta manera podríamos verificar que los fenómenos identificados como desviaciones de la norma monolingüe no existían en la situación anterior al contacto. Así pues, siempre será útil complementar los estudios de cambio aparente mediante la observación directa de estados de lengua antiguos a partir del análisis de documentos históricos producidos en el entorno bilingüe y de este modo trazar en tiempo real la historia de los fenómenos de interferencia.

El objetivo de este trabajo es explorar las posibilidades de la investigación histórica en el estudio del contacto de lenguas a partir del análisis de dos fenómenos de la variedad del español en contacto con el catalán en Mallorca (si bien también tendremos en cuenta la información existente sobre las variedades de castellano de otros territorios de habla catalana). Nuestra intención es mostrar cómo el estudio multidimensional del cambio lingüístico combinando datos y técnicas de análisis sincrónico y diacrónico contribuye a enriquecer nuestro conocimiento del cambio lingüístico en situaciones de contacto de lenguas.

En primer lugar repasaremos los principales mecanismos que explican tanto la aceleración del cambio como el efecto contrario, la desaceleración, en situaciones de contacto de lenguas con especial atención a la situación de Mallorca. A continuación expondremos la metodología que vamos a seguir

para combinar la visión sincrónica y diacrónica en el análisis de fenómenos supuestamente inducidos por contacto en el castellano de Mallorca. Después, en la parte principal del trabajo, estudiaremos dos fenómenos para ilustrar las cuestiones teóricas y metodológicas expuestas: el empleo del *pedir* con usos semánticos y sintácticos propios de *preguntar* y la distribución de los usos del futuro sintético *cantaré* frente al analítico *voy a cantar*.

## 2. EL CONTACTO DE LENGUAS COMO INHIBIDOR DEL CAMBIO LINGÜÍSTICO

Como ya hemos señalado, es un hecho conocido que el contacto de lenguas actúa en numerosas ocasiones como promotor del cambio lingüístico. Por el contrario, son relativamente pocos los estudios que analizan situaciones en las que el contacto de lenguas promueve el conservadurismo lingüístico, es decir, la preservación de variantes que son recesivas en variedades monolingües de la misma lengua. El mecanismo que propicia este resultado es relativamente sencillo: cuando existe en una lengua una tendencia de cambio, es decir, una variante innovadora que está extendiéndose a más contextos y aumentando su frecuencia textual, el avance de esta nueva forma puede verse frenado en una situación de contacto de lenguas debido a que: (i) la estructura innovadora no tiene un paralelo estructural en la lengua en contacto o (ii) la variante recesiva sí tiene una estructura paralela en la lengua en contacto. En este tipo de situaciones el cambio podría verse frenado, y la variante innovadora se extendería con mayor lentitud que en variedades monolingües que se encuentran libres del efecto inhibitor del contacto lingüístico.

Estos factores de tipo estructural pueden verse además reforzados cuando la lengua en cuestión es la variedad socialmente dominante, es decir, que se adquiere y usa en contextos formales. En estos casos, la situación sociolingüística favorecerá la retención de la variante tradicional, sobre todo cuando la forma innovadora es característica del registro coloquial. La razón es que los hablantes que adquieren y emplean la lengua que es socialmente dominante en contextos formales estarán menos expuestos a la variante innovadora y por tanto la emplearán más raramente en sus producciones orales.

Como ejemplos concretos del mecanismo que acabo de describir sirva la evolución de dos fenómenos fonéticos en el español hablado en territorios de habla catalana. Como bien es sabido, el yeísmo es una tendencia en el español peninsular que ha llegado a convertirse en el patrón dominante de

las variedades urbanas. La existencia de una distinción más robusta entre /λ/ y /j/ en el catalán tiene el efecto de frenar este cambio, de tal modo que entre los hablantes cuya primera lengua es el catalán hay índices más altos de distinción que entre los que son dominantes en castellano (Romera 2003: 371-72). Lo mismo se puede decir del debilitamiento de [d] en los participios en -[ado], una tendencia creciente en el castellano que en el caso de los bilingües dominantes en catalán se ve frenada por la tendencia contraria de esta lengua, en la que la estructura equivalente emplea un sonido articulado con una tensión mucho mayor (Blas Arroyo 2007: 269-270; Romera 2003: 373). Además, como la mayor parte de los bilingües que son dominantes en catalán aprenden y usan el español en contextos formales (escuela, medios de comunicación) están menos expuestos al debilitamiento y pérdida de [d] que es característico de los registros coloquiales del español.

Como se detalla en Enrique-Arias (2010) este mecanismo podría explicar la presencia de una serie de fenómenos en el castellano de Mallorca, como la concordancia negativa preverbal (*nadie no ha venido*), las colocaciones con el verbo de apoyo *hacer* (*hacer asco* ‘dar asco’), la preferencia por el empleo de *haber de* como perífrasis de obligación, ciertos usos de las preposiciones *a* y *en*, el empleo del verbo *pedir* con el significado de ‘preguntar’, y la tendencia al empleo de un sistema de dos miembros en la deixis de los demostrativos, entre otros. El análisis de datos históricos muestra que estos rasgos, que son característicos de la variedad de castellano hablada en Mallorca, estaban ya presentes en los textos castellanos producidos por mallorquines en el siglo XVIII. A su vez, estos fenómenos tenían cierta presencia en el castellano general en la época en que esta lengua fue introducida en la isla de Mallorca. La continuidad de estos fenómenos en el castellano de ese territorio sugiere que, al menos en lo que afecta a ciertas estructuras, el contacto con el catalán ha producido la preservación de usos lingüísticos que han retrocedido o incluso desaparecido en el español general, pero que en Mallorca han sido reforzados por la presencia de una estructura paralela en catalán.

En definitiva, no parece que el bilingüismo tenga que ser siempre promotor de cambios lingüísticos sino que también puede promover la tendencia contraria, es decir, la retención de variantes tradicionales. Ahora bien, para poder analizar el origen de estas estructuras y saber si son innovaciones inducidas por el contacto con el catalán o, por el contrario, su origen está en la preservación de rasgos del castellano que estaban presentes en la lengua general en el momento en que se introdujo en la isla, es pertinente hacer un análisis histórico.

El proceso de castellanización cultural y lingüística de Mallorca comienza a partir de la unión dinástica de Castilla y Aragón a finales del siglo xv, con la subsiguiente primacía política, económica y cultural de Castilla entre los reinos peninsulares. Parece ser que ya en el siglo xvi los habitantes instruidos de la ciudad de Palma tenían, como mínimo, competencia pasiva en castellano, como ponen de manifiesto numerosas noticias de impresión de libros y celebración de sermones, poesía de circunstancias y representaciones de teatro en esa lengua. El proceso de castellanización lingüística se vio acentuado durante el siglo xviii con las diferentes iniciativas que los Borbones pusieron en marcha para sustituir al catalán en los ámbitos administrativos, educativos y comerciales (Martinez i Taberner 2000: 355-359). En definitiva, se puede decir que a partir de las últimas décadas del siglo xviii ya hay en Mallorca un sector de la población catalanohablante que emplea el castellano como lengua habitual para usos escriturarios y para comunicarse oralmente con los castellanohablantes residentes o visitantes en la isla (Enrique-Arias 2014: 362). Una de las consecuencias de los cambios que se van produciendo durante esta época es una situación diglósica en la que solamente el castellano es empleado en usos formales de la esfera pública (administración, educación, medios de comunicación). Paralelamente el desconocimiento del castellano se consideraría como propio de individuos pobres e incultos mientras que su dominio sería, además de algo prestigioso, una cualidad indispensable para el progreso social. Esta situación de diglosia continúa hasta la década de 1980, cuando con la instauración de la democracia se recupera el uso del catalán en todos los ámbitos de la vida pública.

En el momento actual el conocimiento del castellano en Baleares es prácticamente universal mientras que el conocimiento del catalán está condicionado por el lugar de origen de los hablantes. Según datos de 2010 recogidos por el Gobierno de las Islas Baleares, entre los individuos nacidos en Baleares 94% declararon que hablan la lengua, un porcentaje que baja al 29% si se considera a los nacidos fuera de las islas. Globalmente el 63,4% de la población de las Baleares es capaz de hablar el catalán o, lo que es lo mismo, es bilingüe en castellano y catalán (Melià 2011)<sup>1</sup>.

De la historia del contacto entre el castellano y el catalán nos interesa destacar dos aspectos: en primer lugar, que estamos ante una situación de bilingüismo de sociedad que se ha prolongado durante varios siglos por lo

---

<sup>1</sup> Los datos proceden de la Encuesta Modular de Hábitos Sociales 2010 elaborada per el Instituto de Estadística de las Islas Baleares; para un análisis detallado de los datos véase Melià (2011).

que se dan las circunstancias para que se produzca transferencia de rasgos entre las dos lenguas; y en segundo, que durante la mayor parte del periodo en que se ha dado este contacto el catalán ha estado en una situación de lengua subordinada al castellano.

### 3. CORPUS Y METODOLOGÍA

Para poder hacer un estudio multidimensional histórico y sincrónico del castellano de Mallorca con una base empírica sólida precisaríamos los siguientes elementos:

- (1) Para el análisis diacrónico:
  - (a) Corpus histórico de textos en castellano producidos por bilingües catalanohablantes. Se trataría de textos de inmediatez comunicativa, como cartas personales o testimonios judiciales en estilo directo.
  - (b) Corpus históricos de control, tanto de español producido en un entorno monolingüe como de catalán de Mallorca; estos corpus de control tendrían que tener características semejantes al corpus histórico de bilingües en lo que respecta a extensión, tipología textual y perfil sociolingüístico de los que escriben.
  
- (2) Para el análisis sincrónico:
  - (a) Corpus de entrevistas recogidas en Mallorca con estratificación sociolingüística distinguiendo, además de los diversos grupos de sexo, edad y nivel de estudios, diferentes grados de bilingüismo (dominantes en catalán, dominantes en castellano, bilingües equilibrados).
  - (b) Acceso a un corpus comparable de entrevistas de español compilado en un entorno monolingüe así como disponibilidad de datos del catalán de Mallorca.
  - (c) Datos de tareas de elicitación para el estudio de fenómenos concretos como juicios de gramaticalidad, tests de comprensión e interpretación de oraciones y tests de producción oral y escrita dirigida elicitada.

En el momento actual no existen fuentes de datos fiables y de fácil acceso con las características que acabamos de señalar. Los corpus diacrónicos convencionales no permiten consultar textos producidos en Mallorca y tampoco hay corpus de entrevistas sociolingüísticas del castellano de la isla. Con el objetivo de superar esta situación el grupo de investigación que coordina el que suscribe en la Universidad de las Islas Baleares lleva varios años desarrollando herramientas metodológicas diseñadas para poder estudiar la diacronía del castellano en Mallorca. En lo que respecta a la vertiente diacrónica el equipo investigador ha comenzado a poner en marcha el *Corpus Mallorca* ([www.corpusmallorca.es](http://www.corpusmallorca.es)), una colección de testimonios judiciales y cartas privadas de los siglos XVIII y XIX escritas en español por catalanohablantes. El proyecto se enmarca en la red internacional CHARTA ([www.charta.es](http://www.charta.es)), formada por grupos de investigación dedicados a la edición electrónica de documentos antiguos, que ha establecido un estándar para su presentación<sup>2</sup>. La vinculación a la red permitirá el acceso a corpus comparables confeccionados en territorios monolingües del español que servirán como corpus de control contrastable con el corpus mallorquín.

En el plano sincrónico tampoco existe a día de hoy un corpus sociolingüístico adecuado para el estudio del castellano de Mallorca. Esta situación cambiará sustancialmente una vez esté disponible el corpus PRESEEA Palma, en fase de transcripción, con entrevistas recogidas con estratificación sociolingüística en la ciudad de Palma. La configuración en red de los diferentes grupos PRESEEA permitirá asimismo contrastar los datos de Mallorca con los de variedades de español monolingüe peninsular<sup>3</sup>.

En el análisis que presentamos en las líneas que siguen, de tipo preliminar y programático, emplearemos los materiales disponibles en este momento. Para el estudio diacrónico utilizaremos el archivo epistolar de Cecilia Zaforteza, que contiene 250 cartas (la mitad aproximadamente en castellano y la otra mitad en catalán) fechadas entre 1739 y 1788 (para una descripción completa cf. Enrique-Arias 2014) complementado con la documentación

---

<sup>2</sup> La red CHARTA (Corpus Hispano y Americano en la Red: Textos Antiguos) es un proyecto destinado a la publicación en internet de un corpus de documentos antiguos de España e Hispanoamérica siguiendo un sistema riguroso de presentación con el fin de satisfacer distintas necesidades de investigadores y usuarios en general (para más información véase la página electrónica del proyecto en [www.charta.es](http://www.charta.es)).

<sup>3</sup> PRESEEA (Proyecto para el Estudio Sociolingüístico del Español de España y América) tiene como uno de sus objetivos principales la creación de un corpus de lengua española hablada representativo del mundo hispánico en su variedad geográfica y social ([www.preseea.linguas.net](http://www.preseea.linguas.net)).



del *Corpus Mallorca*, que incluye en la actualidad transcripciones de una treintena de expedientes con declaraciones judiciales datadas entre 1769 y 1841. Para obtener datos de la variedad actual del castellano en contacto con el catalán en Mallorca utilizaremos las descripciones disponibles (Moll 1961, Serrano Vázquez 1996, Romera 2003) así como ejemplos obtenidos de textos publicados en la prensa local.

Asimismo, para poder contrastar la lengua producida en el entorno bilingüe de Mallorca con el español de Cataluña y de las variedades monolingües peninsulares hemos utilizado una tarea de juicios de gramaticalidad. Se trata de un cuestionario anónimo distribuido en línea entre profesores y alumnos universitarios de Palma, Barcelona, Gerona, Salamanca, Madrid y Sevilla, quienes a su vez lo han redirigido a sus contactos. El cuestionario recoge la siguiente información de los participantes: sexo, nivel de estudios, lugar de residencia y lengua o lenguas nativas (en el caso de los bilingües los participantes tenían que especificar con qué lengua se sienten más cómodos o se identifican más). La Tabla 1 a continuación resume las características de los participantes en el estudio<sup>4</sup>.

	Baleares		Cataluña		Español monolingüe			TOTAL
	L1 Cat.	L1 Cast.	L1 Cat.	L1 Cast.	Castilla y León	Madrid	Andalucía	
Número	65	112	43	44	61	72	47	444
Mujeres	41 (63%)	71 (63%)	27 (63%)	34 (77%)	38 (62%)	43 (59%)	31 (66%)	285 (64%)
Universitarios	47 (72%)	79 (70%)	39 (91%)	39 (88%)	57 (93%)	66 (90%)	42 (89%)	369 (83%)

**Tabla 1.** Composición de participantes en el estudio de juicios de gramaticalidad

Como puede apreciarse en la Tabla 1 la muestra de participantes de cada área geográfica es razonablemente homogénea, con dominio de mujeres (64%) y de individuos con formación universitaria (83%). En total son 444 informantes, de los cuales 180 son monolingües de español que viven en comunidades castellanohablantes y 264 viven en territorios catalanohablantes (177 en Baleares y 87 en Cataluña) con representación tanto de individuos que declaran castellano como su L1 como de participantes que hacen lo propio

<sup>4</sup> El diseño del cuestionario y la recogida de datos se ha hecho en colaboración con Laura Tudurí Cladera (Tudurí Cladera 2014).

con el catalán<sup>5</sup>. La tarea consiste en evaluar una serie de enunciados con estructuras características del castellano de Baleares usando una escala del 1 (“completamente inaceptable”) al 7 (“completamente natural y aceptable”) pasando por el valor intermedio 4 (“aceptable pero suena un poco raro”).

#### 4. TRANSFERENCIA LÉXICO-SEMÁNTICA CON CONSECUENCIAS SINTÁCTICAS

Uno de los resultados típicos del contacto de lenguas es el préstamo léxico con consecuencias sintácticas, como en el fenómeno que describimos a continuación. En el español estándar monolingüe los verbos *pedir* (‘solicitar a alguien que dé o haga algo’) y *preguntar* (‘hacer una pregunta’) tienen, además de significados distintos, comportamientos sintácticos diferentes: solo *preguntar*, pero no *pedir*, puede subcategorizar una cláusula interrogativa, como se muestra en (1a). Por el contrario, el catalán tiene un verbo, *demanar*, que permite ser usado en ambos sentidos y además puede subcategorizar cláusulas interrogativas (1b).

- (1) a. Le \*pidió / preguntó cómo estaba.  
b. Li va demanar com estava.

Aparentemente como resultado del contacto con el catalán, *pedir* en el español hablado en Mallorca tiene las propiedades semánticas y sintácticas del *demanar* catalán. Por consiguiente, en Mallorca y otras áreas catalanohablantes el enunciado (1a) con *pedir* es aceptable y de hecho es muy normal oír expresiones como *pedí por María o me pidió cuánto cuesta el libro* (Moll 1961: 471; Serrano Vázquez 1996: 384-85; Selfa Sastre 2014:129). Los siguientes ejemplos recogidos de la prensa local ilustran este uso<sup>6</sup>:

- (2) a. Nos **pidieron** qué hacíamos aquí. [UH 7/02/06]  
b. Así que llega al convento de Calvià, y **pide** por las langostas. [UH 29/10/06]  
c. Chenoa se sometió ayer a la entrevista [...] “No sé qué **me pedirán**”. [DM 16-05-2008]

<sup>5</sup> No hay un control estricto de quién rellena la encuesta por lo que los participantes podrían introducir respuestas falsas; en cualquier caso tenemos suficiente número de respuestas como para asegurar que cualquier desviación deliberada solo producirá una muy pequeña variación aleatoria; por consiguiente consideramos que los resultados son válidos para la exploración preliminar que pretendemos llevar a cabo en esta investigación.

<sup>6</sup> Identifico los ejemplos de la prensa local de Palma de Mallorca especificando la fuente (UH para *Última Hora* y DM para el *Diario de Mallorca*) y fecha de publicación.

- d. remitió a su declaración efectuada ayer ante los agentes de la Guardia Civil. «**Pídeselo** a ellos» [DM 05/01/2013]

Con el objeto de estudiar el alcance de esta estructura en la población bilingüe catalanohablante y monolingüe de diversas áreas geográficas en el cuestionario se han incluido las siguientes preguntas relevantes con respecto a los usos de *pedir* en el castellano de Baleares:

- (3) a. En el examen me pidieron las capitales de Europa  
 b. Mi vecina se acercó y me pidió por María.  
 c. El policía se acercó y nos pidió qué estábamos haciendo allí  
 d. Sebastián me pidió cuál era el camino hacia el castillo.  
 e. Si no sabes llegar a mi despacho pide en recepción

Las cuestiones presentan diferentes grados de alejamiento de la norma monolingüe castellana. En el enunciado (3a), en que *pedir* va acompañado de un sintagma nominal con función de complemento directo, el verbo es entendido como ‘pedir información’; se trataría de una estructura posible en castellano monolingüe que emplea expresiones como *pedir el nombre* o *pedir la dirección*. La pregunta (3b) presenta un uso de *pedir por* que no es aceptable en el castellano general. Las preguntas (3c-e) en que *pedir* subcategoriza una interrogativa indirecta representan enunciados que tampoco son posibles en castellano monolingüe estándar. La pregunta (3e) también supone una transgresión de las normas del castellano estándar en el que en su uso intransitivo *pedir* se entiende como ‘pedir limosna’ (por ejemplo *Juan se vio obligado a pedir por las calles*) y no ‘preguntar’. La Tabla 2 presenta resumidas las puntuaciones recibidas por los diferentes enunciados.

	Baleares		Cataluña		Español monolingüe
	L1 Cat.	L1 Cast.	L1 Cat.	L1 Cast.	
a pedir + SN	5,58	6,10	5,15	4,88	4,36
b pedir por	4,94	4,54	5,13	2,73	1,85
c pedir qué	5,00	5,05	4,89	3,35	1,81
d pedir cuál	5,19	5,42	5,19	3,55	2,02
e pedir intrans.	4,44	4,81	4,92	3,08	1,91
Media b, c, d y e	4,89	4,95	5,03	3,17	1,89

**Tabla 2.** Gramaticalidad de estructuras con *pedir* por ‘preguntar’

Como ya habíamos adelantado, el primer enunciado (*en el examen me pidieron las capitales de Europa*) tiene un alto grado de aceptabilidad en todos los grupos, tanto entre los participantes de entorno catalanohablante como los de español monolingüe. El hecho de que el verbo *pedir* en ese contexto pueda incluir el significado de ‘preguntar’ da muestra de que uno y otro lexema comparten un espacio semántico común que ha facilitado la extensión de *pedir* hacia contextos propios de ‘preguntar’ que hoy no se aceptan en el español monolingüe. En lo que respecta a las demás preguntas (b, c, d y e) con los usos que son idiosincráticos de la variedad de contacto con el catalán los resultados muestran que en Baleares tienen un alto grado de aceptabilidad tanto entre los hablantes que se consideran L1 catalán (4,89 de media) como los de L1 castellano (4,95); ello indica que se trataría de una estructura completamente integrada en el castellano de Baleares. En Cataluña tenemos una situación diferente, pues en esa región los dominantes en catalán aceptan las estructuras (5,03) mientras los castellanohablantes las sitúan por debajo del nivel de aceptabilidad (3,17) que estaría en torno al 4. Por último los hablantes monolingües de español rechazan claramente estos usos de *pedir* en lugar de *preguntar* (1,89).

Por otro lado, el corpus histórico de documentos producidos en el entorno bilingüe de Mallorca muestra que esta estructura no es nueva en el español de las tierras mallorquinas. Las cartas y los testimonios judiciales del siglo XVIII y XIX contienen ejemplos en los que *pedir* subcategoriza una cláusula interrogativa:<sup>7</sup>

- (4) a. Margarita Dezclapés me ha embiado recado por que passasse a verla; y ha sido por **pedirme**, de Juana, si es de buena gente, si es de mala condicion, si enfermiza, si sirve bien [ZA, 159, Palma, 1780]
- b. le acometieron y pegaron primero sin mas fundamento que el haver antes el declarante llamado aparte a Visente Palou y **pedidole** si contenia verdad que havia dicho bruta a su muger [AA 974/10, 30v, Buñola, 1833]

Una vez establecido que estos empleos que se alejan de la norma monolingüe castellana tienen cierta trayectoria en el castellano de Baleares no nos queda sino comprobar si también existían en el castellano general

<sup>7</sup> Identifico los ejemplos del corpus histórico señalando la fuente y localización (ZA para el epistolario de Cecilia Zaforteza, con indicación del número de carta, localidad y el año, y AA para los documentos del Archivo de la Audiencia seguido de la signatura del expediente, número de folio, lugar y año).

antes del contacto con el catalán, un aspecto que no se ha tenido en cuenta en los estudios llevados a cabo desde una perspectiva puramente sincrónica. En este sentido es relevante señalar que el uso de *pedir* en el sentido de ‘hacer una pregunta’ no constituye ninguna rareza a lo largo de la historia del español. Corominas y Pascual (sv. *pedir*) afirman que “antiguamente se había empleado *pedir* con el valor de *preguntar* [...] como ocurre con el fr[ancés] *demander*, cat[alán] *demanar*”. En el *Diccionario de autoridades* y demás diccionarios académicos desde la primera edición de 1737 hasta la de 1852 *pedir* se define en su segunda acepción como “preguntar, o informarse de otro de alguna cosa”. Además, el ejemplo aportado por *Autoridades* para ilustrar el uso correcto de este verbo es *pedían cuál era el camino de los perdidos*, donde *pedir* aparece subcategorizando una cláusula interrogativa (Real Academia Española 1961: s.v. *pedir*). El ejemplo procede de *El criticón* de Baltasar Gracián (Correa Calderón 1971: vol. I p. 69), obra publicada en 1651 en la que aparece otro caso: *Llamáronlos de otra tienda [...] y pidiendo el valor, dijeron...* (vol. I, p. 218).

Por su parte Rufino José Cuervo en su *Diccionario de construcción y régimen de la lengua castellana* (1994 s.v. *pedir*) también se hace eco de este uso supuestamente anticuado que ilustra con numerosos ejemplos históricos de entre los que selecciono los siguientes:

- (5) a. Pues **pidiendo** Diego García de Paredes por el capitán general y otros capitanes para que allí se juntasen les dijo... [*Crónicas del Gran Capitán*, 1554]
- b. **Pedí** por la cuenta. Mi compañero dijo que la dejase, que él daría recaudo [Mateo Alemán, *Guzmán de Alfarache*, 1599]
- c. Echa los nombres en una urna, saca uno por suerte, **pide** al pueblo lo que se ha de hacer dél; y todos le condenan a muerte [Diego de Saavedra Fajardo, *Empresas Políticas*, 1640]

También es posible encontrar tres ejemplos en el *Quijote* de Avellaneda, impreso en Tarragona en 1614: **pidiéndola** de su salud (Riquer 1972: vol. 2, 91), **pidió** por el predicador (vol. 2, 143), **pidiendo** [...] el nombre de la priora de tal convento (vol. 2, 146).

Además de las obras lexicográficas y estudios de tipo histórico son también varios los diccionarios del español actual que recogen el uso de *pedir* en su acepción de ‘interrogar, preguntar’ con etiquetas especiales. El

diccionario de la Real Academia Española (1992) lo considera “desusado”; Moliner (1987) lo etiqueta como “no usual” y Seco (2011) recoge la definición ‘preguntar por alguien’ que tilda de “regional” e ilustra con dos ejemplos sacados de la obra del escritor barcelonés Manuel Vázquez Montalbán.

Respecto de la adscripción geográfica del fenómeno que nos ocupa, Martín de Riquer afirma que el uso de *pedir* por ‘preguntar’ es aragonesismo de los escritores del XVII que sigue vigente en el habla actual (Riquer 1972 vol. 1, lxv). Ya hemos visto cómo en el *Quijote* de Avellaneda, de autor supuestamente aragonés, hay varios ejemplos de *pedir* ‘preguntar’ junto a otros aragonesismos, y lo mismo sucede con Gracián (1601-1658), que era aragonés de Calatayud. Asimismo cabe señalar que Gonzalo Correas en su *Vocabulario de refranes y frases proverbiales* de 1627 nos da la escueta pero esclarecedora indicación “Pedir, por preguntar. En Aragón” (cf. Luis Combet *et al.* 2000: 1038).

Ahora bien, aunque el fenómeno sea característico de Aragón o presente mayor intensidad en esa área geográfica, parece que en época antigua tenía mayor extensión como indican los ejemplos aducidos por Cuervo sacados de textos escritos por un autor murciano (Diego de Saavedra Fajardo) y otro sevillano (Mateo Alemán). Además, el hecho de que el diccionario de *Autoridades* recoja este uso indica que en la época de implantación del castellano en Baleares (siglos XVIII y XIX) era una estructura que los hablantes instruidos castellanos consideraban aceptable.

En definitiva, el análisis que acabamos de llevar a cabo nos permite concluir que los usos de *pedir* en el español hablado en Mallorca no parecen ser meramente el resultado de un préstamo *ex novo* de estructuras sintácticas o reglas del catalán al español. El contacto con el catalán no ha resultado tanto en la introducción de una estructura innovadora, sino en la preservación de un uso lingüístico que existía en el español monolingüe, reforzado por la existencia de una estructura paralela en catalán.

## 5. VARIACIÓN Y CAMBIO EN LA EXPRESIÓN DEL FUTURO

### 5.1. Diferencias y paralelos entre castellano y catalán

El catalán y el castellano son lenguas muy próximas estructuralmente, tanto por el origen latino que comparten como por la convivencia prolongada que han mantenido a través del tiempo, especialmente en los territorios

donde, como es el caso de Mallorca, se da un alto grado de bilingüismo de sociedad. Como no podía ser de otro modo, esta circunstancia ha facilitado el que se den ciertos paralelos estructurales si bien también encontramos diferencias.

En lo que concierne a la expresión del futuro, el castellano y el catalán exhiben una primera asimetría evidente. Si por un lado las dos lenguas tienen en común el futuro morfológico romance derivado de la perífrasis latina *infinitivo* + HABERE (AMARE HABEO > *amar-é*), el español ha desarrollado otra forma de futuro a partir de la perífrasis con *ir* (*voy a amar*) que no tiene un paralelo estructural en catalán; bien al contrario, el equivalente literal en catalán, la perífrasis con *anar*, tiene un significado de pretérito (*vaig amar* ‘amé’).

La distribución de los dos futuros en el sistema verbal del español ha experimentado un cambio desde el siglo XVII por el cual el futuro morfológico *amaré* ha cedido terreno progresivamente al futuro perifrástico *voy a amar* (Aaron 2010). Son abundantes los estudios que demuestran este avance, que ha tenido especial desarrollo en las variedades americanas (Silva-Corvalán y Terrell 1989; Sedano 2007). Incluso se ha postulado que en el español del suroeste de los Estados Unidos el contacto con el inglés ha favorecido aún más el avance del futuro perifrástico (Gutiérrez 1995; 2002). Sin embargo, en las zonas donde el español coexiste con el catalán este proceso de sustitución está menos avanzado (Blas Arroyo 2004: 1068). Esta ralentización de una tendencia de cambio que es prácticamente panhispánica se explicaría por la asimetría estructural antes explicada: la tendencia histórica a remplazar el futuro morfológico por el perifrástico con *ir* se ve contrarrestada en los bilingües en catalán por el hecho de que, mientras la variante tradicional tiene un paralelo en el futuro catalán *amaré*, la forma perifrástica innovadora no tiene un equivalente en catalán. Se trata por tanto de un caso en el que el contacto de lenguas tiene como efecto la inhibición de un cambio lingüístico (Enrique-Arias 2010: 104-105). Blas Arroyo (2007) ha demostrado este efecto inhibitorio del contacto lingüístico en el empleo de los futuros en el español de Castellón. En su estudio, en el que analiza datos de hablantes bilingües, los individuos con mayor dominio del catalán presentan sistemáticamente un mayor porcentaje de uso del futuro morfológico en comparación con los hablantes dominantes en castellano, es decir, el contacto de lenguas en esta situación promueve la ralentización de una tendencia de cambio del español general.

Otra diferencia notable entre castellano y catalán en lo que respecta a los usos del futuro sintético es que en castellano esta forma ha adquirido



significados modales de tipo epistémico, en particular un valor de conjetura que en el catalán tradicional no existe. El equivalente catalán de un enunciado como *ahora serán las doce* se construye con una perífrasis del verbo *deure* ‘deber’ y el infinitivo: *ara deuen ser les dotze*. Se trata por tanto de un contexto en que en principio, y al contrario de lo que sucede en la referencia temporal, el castellano utiliza el futuro morfológico en más contextos que el catalán. No obstante no son infrecuentes en catalán usos del futuro con valor de conjetura del tipo *ara seran les dotze* (supuestamente por influencia castellana), usos que son extraños al catalán tradicional (Badia i Margarit 1985: 391). Asimismo, cabe notar que en la documentación histórica de español producido en el entorno bilingüe de Mallorca el uso epistémico del futuro simple aparece con cierta frecuencia. En no pocas ocasiones encontramos ejemplos en escritos de bilingües que, a pesar de acusar una fuerte interferencia del catalán con respecto a otros fenómenos, han adquirido este uso verbal que es en principio extraño al catalán.

- (6) a. y dize dicho declarante que dicho muchacho **será** de edad de 13 a 15 años (AA 950/3 8v 1769)  
 b. es hombre de edad de 23 a 25 años [...] **será** de cerca de ocho palmos alto (AA 950/3 9r 1769)  
 c. que **habrá** unos tres años que a cosa de las diez de la noche lo encontró (AA 974/5 2v 1833)

El castellano y el catalán presentan diferencias también en cuanto a los contextos sintácticos en los que puede aparecer el futuro sintético. En castellano actual el futuro morfológico no se considera gramatical en una serie de subordinadas en las que el catalán tradicional emplea esa forma. Una de estas estructuras es la prótasis de la condicional donde supuestamente no se podrían dar enunciados del tipo *\*avísanos si vendrás a la fiesta* (lo gramatical sería *avísanos si vienes / vas a venir a la fiesta*) mientras que en catalán sí que es posible este uso: *avisa'ns si vindràs a la festa*. Asimismo el catalán permite el empleo del futuro en subordinadas temporales y modales así como en sustantivas de relativo, usos en los que el castellano requiere el empleo del subjuntivo (cf. Badia i Margarit 1985: 430): *saldrás cuando termines* frente a *sortirás quan acabaràs; el que retroceda será castigado* tiene como equivalente *el que retrocedirà serà castigat*. Ahora bien, estas construcciones canónicas del catalán conviven con la variante con subjuntivo (*sortirás quan acabis; el que retrocedeixi serà castigat*) lo cual sería una manifestación de un proceso antiguo de expansión del subjuntivo en estas



construcciones catalanas que en todo caso estaría ayudado por la existencia de una construcción paralela castellana. Y al mismo tiempo que se da esa influencia del castellano en el catalán se puede oír a catalanohablantes transfiriendo la estructura con futuro en los enunciados del castellano que no lo permitirían en la variedades monolingües: *cuando volverán los recibiremos* (Colón 1967: 203).

Con el objeto de contrastar el grado de aceptabilidad del uso del futuro en contextos sintácticos característicos del catalán (subordinadas condicionales, sustantivas y temporales) hemos incluido en el cuestionario las siguientes oraciones:

- (7) a. Avísame si vendrás a la fiesta  
 b. El caso es muy difícil pero haremos lo que podremos  
 c. Añade los fideos cuando el caldo estará bien caliente.

	Balears		Cataluña		Español monolingüe
	L1 Cat.	L1 Cast.	L1 Cat.	L1 Cast.	
a) condicional	4,97	4,88	4,19	4,7	3,84
b) sustantiva	3,25	2,27	3,92	2,67	1,48
c) temporal	1,98	1,54	1,79	1,60	1,51
Media	3,4	2,89	3,3	2,99	2,27

**Tabla 3.** Gramaticalidad de usos de futuro en subordinadas

Los datos de la Tabla 3, donde se resumen los resultados de la encuesta en lo que respecta a los enunciados (7a-c), muestran efectos interesantes de la transferencia lingüística del catalán en la gramática de los bilingües. Como era de esperar, en términos globales todas las estructuras tienen un grado más alto de aceptación en el ámbito bilingüe que en el monolingüe, y dentro de este último son siempre los hablantes con L1 catalán los que otorgan una aceptabilidad más alta. Por otro lado no todos los enunciados tienen el mismo grado de aceptación: el uso del futuro en la condicional es el mejor aceptado, y de hecho, entre los hablantes monolingües *avísame si vendrás a la fiesta* queda muy cerca del 4, que marcaría el límite de la aceptabilidad (3,84). El empleo del futuro en la sustantiva *haremos lo que podremos* nos da un contraste nítido entre los bilingües con L1 catalán en Cataluña (3,92) y el rechazo claro de los participantes de áreas monolingües de castellano

(1,48). Por último, el enunciado menos aceptado es el que emplea futuro en la temporal (*cuando el caldo estará bien caliente*), en donde apenas se aprecia variación entre los diferentes grupos de participantes (la valoración más alta, la del 1,98 del grupo L1 catalán de Mallorca no dista mucho del 1,51 que otorgan los monolingües). Sin embargo es precisamente este uso el que vemos reflejado con cierta frecuencia en la documentación antigua: en la fórmula del juramento que precede a las declaraciones judiciales del corpus histórico encontramos futuro allí donde documentos semejantes del español monolingüe emplean subjuntivo (*decir verdad de lo que supiere y fuere preguntado*).

- (8) a. prometió decir verdad a la regia corte de lo que **será** sabidor y **será** interrogado (AA 950/3 2r, Palma, 1769)  
 b. prometió decir verdad de todo lo que **sabrá** y **será** preguntada (AA 950/13, 1v, Esporlas, 1778)

Como hemos visto en este rápido análisis contrastivo de los usos del futuro en castellano y catalán hay toda una serie de estructuras susceptibles de transferirse entre una y otra lengua (frecuencia relativa de futuro sintético y perifrástico, valor epistémico del futuro simple, alternancia de subjuntivo y futuro en determinados contextos sintácticos). Se trata de fenómenos de interés teórico con repercusiones tanto para el estudio sincrónico como para el histórico pero que hasta la fecha no han sido estudiados en profundidad.

## 5.2. Evolución histórica

Después de presentar algunos de los fenómenos de transferencia que atañen a la distribución de los futuros en castellano y catalán nos detendremos en la cuestión de la expansión del futuro simple a expensas del perifrástico. Como ya hemos adelantado, la evolución de este fenómeno en el castellano hablado en territorios de habla catalana parece indicar un ejemplo de retención de una variable tradicional en una situación de contacto. Así al menos lo indican los datos de las entrevistas sociolingüísticas de Blas Arroyo (2007) en Castellón. Para poder observar cómo funciona este mecanismo desde una perspectiva histórica sería necesario contrastar la evolución de la distribución de las dos formas de expresar el futuro y observar si los documentos en castellano producidos en el entorno bilingüe de Mallorca exhiben una extensión más lenta de la forma innovadora, es decir, del futuro perifrástico formado a partir de *ir a* más infinitivo. Aunque no existen en el

momento actual corpus históricos que nos permitan hacer esta comparación con una base empírica sólida procederemos a hacer un análisis preliminar de la cuestión a partir de los materiales disponibles.

El estudio histórico más completo sobre la evolución de las formas de futuro en castellano es el de Aaron (2010). En su trabajo selecciona varios cortes sincrónicos: a) principios del xvii, b) finales del xviii-principios del xix y c) finales del xx-principios del xxi (la autora considera además el español medieval pero no lo incluimos en la discusión por presentar un número insignificante de ejemplos de futuro perifrástico). En lo que respecta a los resultados, su corpus registra 4% de futuro perifrástico en los siglos xvii y xviii combinados, 13% en el siglo xix y 25% y 59% para, respectivamente, los materiales escritos y orales del periodo más reciente (Aaron 2010: 5). La comparación con los datos disponibles para el español de los bilingües mallorquines sugiere que, efectivamente, en la variedad de contacto se da una evolución más lenta del fenómeno. En las cartas en castellano del archivo epistolar de Cecilia Zaforteza (siglo xviii) encontramos 210 ejemplos de futuro morfológico frente a solo dos ejemplos de futuro perifrástico con *ir*<sup>8</sup>. De igual modo, en la parte documental del corpus la presencia del futuro perifrástico es muy reducida: entre los 132 casos de expresión del futuro para el siglo xviii no tenemos ningún ejemplo del futuro perifrástico, y en el xix solamente hay dos ejemplos frente a 166 de formas de futuro sintético (es decir, el porcentaje de futuro perifrástico no llega al 1%). En la Tabla 4 presentamos de forma resumida una comparación de ambos corpus:

Siglo	Español monolingüe (Aron 2010)	Español bilingüe
Siglos xvii-xviii	4,3% (59/1367)	0,07% (2/280)
Siglo xix	13,1% (77/588)	1,2% (2/168)
Totales	6,9% (136/1955)	0,09% (4/448)

**Tabla 4.** Evolución histórica de la expresión del futuro en áreas bilingües y monolingües

<sup>8</sup> El cómputo está hecho considerando todas las ocurrencias en el corpus epistolar. Si consideramos solo los textos escritos por personas de las cuales tenemos indicios de que son mallorquines bilingües entonces tenemos 148 ejemplos totales (cf. Enrique-Arias 2010: 105, 2012: 18) lo cual continúa siendo un porcentaje muy pequeño de futuro perifrástico (1,3%, o 2/148) y en cualquier caso menor que el de los datos de español monolingüe de Aaron (2010) para el siglo xviii.

La comparación de los datos globales de cada corpus parecen confirmar la hipótesis de que el contacto con el catalán habría tenido el efecto de ralentizar el desarrollo del futuro perifrástico en el castellano de los bilingües, pues tanto los datos de los siglos xvii-xviii como los del xix del corpus peninsular de castellano monolingüe arrojan porcentajes muy por encima de los observados en el corpus mallorquín. Ahora bien, hay que tener en cuenta que existen problemas de comparabilidad entre los dos corpus (cf. Enrique-Arias 2012). En primer lugar hay discrepancias en lo que se refiere a la datación de los textos seleccionados: el corte del corpus peninsular que engloba textos de finales del xviii y principios del xix coincide bastante bien con las fechas del corpus mallorquín pero la autora decidió juntar los datos del xvii y el xviii, lo cual afecta a la comparabilidad de los datos. Los del xix corresponden a dos obras compuestas en la década de 1830, en época cercana al grueso de la parte documental del corpus, pero ya lejanas de la porción epistolar, que abarca los años 1739-1788. Las obras del xviii pertenecen a las dos últimas décadas de la centuria, es decir, coinciden con gran parte de los documentos de la Audiencia de Palma, pero al haberse computado los resultados junto con las obras del xvii se pierde la posibilidad de comparar directamente los datos con los del corpus mallorquín. Para poder observar mejor la evolución histórica del fenómeno estudiado sería además deseable ampliar el arco cronológico hacia la segunda mitad del xix, que no aparece representada en los textos analizados.

La distribución de géneros textuales en los corpus también compromete la comparabilidad. Con la intención de obtener muestras de lengua cercana a las interacciones del discurso oral el corpus usado en Aaron (2010) está compuesto principalmente por obras de teatro (un objetivo que se ve comprometido en el caso de las obras del xvii y una de las del xviii por estar escritas en verso). El corpus mallorquín reúne, por un lado, cartas personales, género que tiende a acercarse más a los registros orales (Biber 1995: 283-300), y por otro, expedientes judiciales, en los que encontramos junto al lenguaje formulístico propio del registro jurídico pasajes con mayor espontaneidad, como las declaraciones de testigos. De todos modos, es interesante recalcar que incluso aquellos empleos más formulísticos no están exentos de la influencia del contacto como ya hemos visto respecto del empleo del futuro en lugar del esperable subjuntivo en la forma del juramento de los testigos (*decir verdad de lo que sabrá*). Asimismo, hay que tener en cuenta que en textos jurídicos y administrativos es frecuente el empleo del futuro sintético para dar instrucciones en usos en los que no hay posibilidad de alternancia con el perifrástico:

- (9) a. así mismo, los que resultaren reos les **embargará** sus bienes y muebles [...] **recibirá** así mismo sumaria informacion de la vida fama y custumbres de los que resultaren reos (AA-950/3, 1v, Palma, 1769)
- b. y de quedar enterado de ello me lo **certificaréis** al pie de la presente (AA-979/14, 38r, Bañalbufar, 1803)

Estos usos fuertemente condicionados por las convenciones del registro jurídico alternan con declaraciones de los testigos en las que hay discurso reproducido en estilo directo, donde encontramos lengua más próxima al registro oral:

- (10) Y respondió Mayol: “no te cases con Ana porque si te casas **comerás** xulla toda tu vida” (AA-979/14, 30v, Bañalbufar, 1803).

Por último, en lo que atañe a los distintos matices de significado de las formas de futuro, cabe mencionar que nos hemos limitado a contabilizar todas las ocurrencias de las formas de futuro estudiadas en uno y otro corpus sin distinguir su función ni su significado. Para hacer una comparación más detallada sería necesario considerar la proporción de los diferentes usos de las formas de futuro (referencia futura, mandato, valor epistémico, etc.) en cada uno de los corpus estudiados. En el momento actual el número tan reducido de ocurrencias de futuro perifrástico en el corpus mallorquín hace del todo imposible un estudio de coocurrencias de los diferentes valores del futuro semejante al llevado a cabo por Aaron (2010).

En lo que respecta a los valores de los cuatro ejemplos de futuro perifrástico contenidos en el corpus (11a-d), se trata en todos los casos de referencia temporal de un futuro en que hay gran proximidad al momento de la enunciación.

- (11) a. En esta **voy a dar** relación a V.M. con toda claridad y pura verdad de hombre de bien el modo de proseguir nuestro viaje (ZA, 125, Nápoles, 1777)
- b. Perico ya está bueno pues luego le **van a dar** el Socorro (ZA, 212, ca. 1778)
- c. Al parecer **va a armarse** un encuentro y competencia entre las dos justicias de Esporlas y Soller que parece del cabo evitar (AA 979/14, 8r, Palma, 1804)

- d. Vuestra Magnánima, administrando justicia, se ha de servir absolver a mi principal [...] por los motivos y razones que **voy a exponer** (AA 974/5, 25r, Buñola, 1834)

En el caso de (11a) y (11c) *voy a dar relación* y *voy a exponer* se refieren a una acción que el hablante está a punto de acometer, es decir, tienen un valor aspectual incoativo que expresa una fase preparatoria de la acción verbal, uso que no es desconocido en el catalán (*anem a veure* ‘vamos a ver’). En (11b) la inminencia de la acción del verbo está realzada con el adverbio *luego* que en el español del XVIII significa ‘enseguida’. Y en (11c) se habla de la inminencia de un suceso que se quiere evitar.

En resumen, el somero análisis de datos históricos que acabamos de hacer ha ilustrado cómo el uso y distribución de los dos futuros en el castellano producido en Mallorca habría sido influido por el contacto con el catalán: los datos disponibles apuntan hacia el empleo menos frecuente del futuro perifrástico y siempre con usos restringidos al ámbito temporal cercano al momento de la enunciación.

## 6. CONCLUSIÓN

Las consideraciones de tipo histórico que hemos ido presentando a lo largo de este capítulo han servido para ilustrar la relevancia de una perspectiva diacrónica en el estudio de fenómenos lingüísticos inducidos por el contacto de lenguas. Somos conscientes de que en el momento actual no tenemos las bases empíricas adecuadas para acometer con garantías una investigación diacrónica del castellano de los bilingües catalanohablantes en Mallorca; con todo, en estas exploraciones hechas de forma preliminar ha quedado claro cómo algunos fenómenos que desde una punto de vista puramente sincrónico se entienden como innovaciones inducidas por la influencia del catalán —el uso de *pedir* con los significados y la sintaxis de *preguntar*, o la preponderancia del futuro sintético frente al perifrástico— son en realidad estructuras que ya estaban presentes en los escritos de los catalonahablantes bilingües en el siglo XVIII. Al mismo tiempo, los fenómenos en cuestión tenían cierta existencia en variedades monolingües del español en el momento en el que la lengua se introdujo en la isla. Este hecho por sí debilita enormemente la visión de que estas estructuras se deben sin más

a la transferencia directa del catalán; más bien se trata de la preservación de estructuras del castellano del siglo XVIII ayudada por la presencia de un paralelo estructural en el catalán. El estudio histórico es pues una vía fructífera para llegar a un mejor conocimiento de cómo se forma una variedad de contacto.

Por otro lado hemos destacado una serie de problemas metodológicos: la escasez de materiales históricos de castellano producido en un contexto bilingüe que sean de fácil acceso a los investigadores, la falta de descripciones detalladas de algunos fenómenos de relevancia teórica y los problemas de comparabilidad entre los datos procedentes de diferentes corpus por contener textos que pertenecen a épocas y registros diferentes. Es sobre todo necesario establecer junto al corpus de castellano producido en el entorno bilingüe de Mallorca un corpus de control de español monolingüe de semejante composición y características. Solamente de este modo podremos avanzar con las deseables garantías en el estudio histórico de los factores sociales, cognitivos y lingüísticos que influyen en la variación y el cambio lingüísticos en situaciones de contacto.

## REFERENCES

- Aaron, Jessi E. (2010): "Pushing the envelope: Looking beyond the variable context". *Language variation and change* 22, pp. 1-36.  
<http://dx.doi.org/10.1017/S0954394509990226>
- Biber, Douglas (1995): *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press
- Blas Arroyo, José Luis (2004): "El español actual en las comunidades del ámbito lingüístico catalán" en Rafael Cano Aguilar (ed.) *Historia de la lengua española*. Barcelona: Ariel, pp. 1065-1086.
- (2007): "El contacto de lenguas como factor de retención en procesos de variación y cambio lingüístico. Datos sobre el español en una comunidad bilingüe peninsular". *Spanish in Context* 4 (2), pp. 263-29.  
<http://dx.doi.org/10.1075/sic.4.2.07bla>
- Colón, Germán (1967): "Catalanismos", en Manuel Alvar *et al.* (coord.): *Enciclopedia lingüística hispánica*. Madrid: Centro Superior de Investigaciones Científicas II, 193-238.
- Combet, Luis, Robert James y Maïte Mir-Andreu (eds.) (2000): Gonzalo Correas. *Vocabulario de refranes y frases proverbiales* (1627). Madrid: Castalia.
- Correa Calderón, Evaristo (ed.) (1971): Baltasar Gracián. *El Criticón*. Madrid: Espasa-Calpe.
- Cuervo, Rufino J. (1994): *Diccionario de construcción y régimen de la lengua castellana*. Santa Fe de Bogotá: Instituto Caro y Cuervo.
- Enrique-Arias, Andrés (2010): "On language contact as an inhibitor of language change: the Spanish of Catalan bilinguals in Majorca", en Anne Breitbarth, Christopher Lucas, Sheila Watts y David Willis (eds.): *Continuity and change in grammar*. Filadelfia / Amsterdam: John Benjamins, pp. 97-118.



- (2012): “Retos del estudio sociohistórico del contacto de lenguas a través de un corpus documental. El caso del castellano en contacto con el catalán en Mallorca”. *Revista de investigación lingüística* 15, pp. 23-46.
  - (2014): “Lengua y escritura en la Mallorca del XVIII: el archivo epistolar de Cecilia Zaforteza”, en Antonio Castillo Gómez y Verónica Sierra Blas (eds.): *Cinco siglos de cartas. Historia y prácticas epistolares en las épocas moderna y contemporánea*. Huelva: Universidad, pp. 313-328.
- Gutiérrez, Manuel (1995): “On the future of the future tense in the Spanish of the Southwest”, en Carmen Silva-Corvalán (ed.): *Spanish in Four Continents: Studies in Language Contact and Bilingualism*. Washington DC: Georgetown University Press, pp. 214-226.
- (2002): “Innovación en el español de los Estados Unidos: ¿Progreso o decadencia?” *Revista de Lingüística Teórica y Aplicada* 40, pp. 113-128.
- Melià, Joan (2011): “La llengua catalana a les Balears. Algunes dades recents sobre usos i coneixements”. *Lluc. Revista de cultura i d'idees* 878, pp. 3-7.
- Martínez i Taberner, Cristina (2000): *La llengua catalana a Mallorca al segle XVIII i primer terç del XIX*. Barcelona / Palma: Publicacions de l'Abadia de Montserrat / Universitat de les Illes Balears.
- Moliner, María (1987): *Diccionario de uso del español*. Madrid: Gredos.
- Moll, Francisco de B. (1961): “El castellano en Mallorca”, en *Studia Philologica. Homenaje ofrecido a Dámaso Alonso por sus amigos y discípulos con ocasión de su 60º aniversario*. Madrid: Gredos, Vol. 2, pp. 469-474.
- Riquer, Martín de (ed.) (1972): Alonso Fernández de Avellaneda. *Don Quijote de la Mancha*. Madrid: Espasa-Calpe.
- Real Academia Española (1963). *Diccionario de autoridades*. Edición facsímil. Madrid: Gredos.
- (1992). *Diccionario de la lengua española*. Madrid: Espasa-Calpe.
- Seco, Manuel (2011). *Diccionario del español actual*. Madrid: Aguilar.
- Sedano, Mercedes (2007). “Future tense expressions in several Spanish corpora”, en Giovanni Parodi (ed.): *Working with Spanish Corpora Research in Corpus and Discourse*. Londres: Continuum, pp. 132-144.

- Romera, Magdalena (2003): “La variedad del castellano actual en Baleares”, *Moenia* 9, pp. 359-381.
- Serrano Vázquez, María del Carmen (1996): “Interferencias léxicas y semánticas en una situación de contacto entre dos lenguas, catalán y castellano”, en Mercé Pujol y Fermín Sierra (eds.): *Las lenguas en la Europa Comunitaria II*. Amsterdam: Diálogos Hispánicos, pp. 375-394.
- Thomason, Sarah (2001): *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press.
- Tudurí Cladera, Laura (2014): *Un fenómeno de transferencia en la gramática bilingüe: pedir en el castellano de Mallorca*. Trabajo de fin de grado. Universitat de les Illes Balears.
- Silva-Corvalán, Carmen (1994): *Language contact and change. Spanish in Los Angeles*. Oxford: Clarendon.
- y Tracy Terrell (1991): “Notas sobre la expresión de futuridad en el español del Caribe”, en Elisabeth Luna Traill (ed.): *Scripta philologica: in honorem Juan M. Lope Blanch*. México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas, Vol. 2, pp. 757-772.



## **II**

# **Spanish in the United States**



## La subida de clíticos en tiempo real y aparente en el español de Houston

Manuel J. Gutiérrez  
*University of Houston*

**Abstract:** A number of Spanish verbal phrases allow speakers to place clitic forms before the first conjugated form or after the last verbal form. The first situation is known as clitic climbing since the pronoun is placed next to the verbal form that is found in the highest position away from its original position (= cliticized to the non-conjugated verb). This is a variationist study that intends to offer information to help to explain the tendencies that are taking place within the general tendency that favors clitic climbing, which has been occurring for the last several centuries. It examines data from the Mexican-American community of Houston, Texas, and the speakers are part of three different generations according to the years in contact with English. The results indicate that gerund verbal phrases favor clitic climbing almost categorically while those with infinitives offer a continuum from those forms that highly favor the position before the verb that expresses the grammatical categories to others that keep clitic forms next to the verb that expresses the lexical meaning of the periphrasis. The use of two sets of data separated by twenty years allows studying the tendencies also using real time.

**Keywords:** Clitic climbing, linguistic variation, linguistic change, bilingualism.



## 1. INTRODUCCIÓN

El sistema gramatical del español contiene entre sus estructuras un grupo de perífrasis verbales que permiten al hablante ubicar los pronombres clíticos después o antes de éstas; algunas se construyen con un verbo léxico en infinitivo mientras otras tienen este mismo componente de la perífrasis en gerundio. El hecho de que los clíticos aparezcan en una posición variable en este contexto ha despertado interés en los especialistas por explorar los factores que podrían determinar la elección que hacen los hablantes (Suñer 1980, Gutiérrez y Silva-Corvalán 1993, Silva Corvalán y Gutiérrez 1995, Davies 1995, Gutiérrez 2008a, 2008b, 2008c, 2010, Aijón Oliva 2010, Troya y Pérez 2011). El fenómeno ha sido discutido desde distintos enfoques; en el presente caso se estudia dentro de un marco variacionista y se intenta determinar algunos factores tanto lingüísticos como sociales que explican en cierta medida el comportamiento de los hablantes. Se estudia el fenómeno con datos producidos por hablantes en contextos conversacionales reales en los que los clíticos aparecen en una u otra posición. El objetivo es, por lo tanto, descubrir los factores que podrían estar gobernando la posición que adoptan estos elementos.

El pronombre clítico puede aparecer después del último elemento de la perífrasis verbal, que corresponde a un verbo en infinitivo o gerundio y que aporta el significado léxico a la construcción, o antes del verbo conjugado, elemento que aporta el significado gramatical. Se habla de *subida* en la segunda situación, ya que el clítico no permanece junto al verbo de la oración incrustada, sino que pasa junto al que se encuentra en una posición más elevada, como en (1) y (2).

- (1) a. Podríamos comprar**lo** / b. **Lo** podríamos comprar.
- (2) a. Están regalándoles caramelos / b. **Les** están regalando caramelos.

En (1) a., el pronombre *lo* se encuentra adherido al verbo *comprar* que aporta el significado léxico a la perífrasis, la acción de comprar, y éste corresponde al verbo principal de la perífrasis *Podríamos comprarlo*. En (1) b., en cambio, el mismo pronombre *lo* sube hacia una posición superior en la estructura de la frase verbal, pues se encuentra antes de *podríamos*, verbo que funciona como auxiliar aportando las categorías gramaticales correspondientes de tiempo, modo/aspecto, persona y número. El proceso descrito es válido para (2) a-b, caso en que la perífrasis se construye con un verbo léxico en gerundio y no en infinitivo como en (1) a-b.



La investigación realizada sobre la posición de los clíticos ha dejado en evidencia distintas tendencias a lo largo de los siglos y definido distintos factores que inciden en alguna medida en la ubicación de estos elementos. Hoy contamos con algunos resultados que consideran información desde el siglo XIII en adelante (Keniston 1937, Luna Traill y Parodi 1974, Torres Cacoullós 1999, Berta 1999, Nieuwenhuijsen 2006). Keniston (1937:89-112) estudia diversos aspectos relacionados con los clíticos del español en la prosa del siglo XVI. Al discutir la posición adoptada por los clíticos durante la mencionada época, el autor establece el siguiente principio general: “In general, the usage of the sixteenth century can be best explained by saying that the object pronouns are enclitic forms. Although they are not separated from the verb, they precede the verb only when they are themselves preceded by some other stressed element in the same breath-group” (p. 89). Para Keniston, este principio gobierna todos los contextos en que aparecen los clíticos y explica el hecho de que éstos ocurran con tanta frecuencia pospuestos al verbo. Lo mismo es válido para la posición de estos elementos en las construcciones compuestas por un auxiliar y otra forma no finita (p. 91). En el estudio de Nieuwenhuijsen (2006), éste comenta resultados de otras investigaciones que apuntan al hecho de que con las formas no finitas se ha observado una tendencia hacia la anteposición de los clíticos, la cual se aprecia desde el siglo XIII. Sin embargo, apunta este autor, esta tendencia pierde vitalidad en el siglo XVI y el proceso inverso comienza a manifestarse durante este período. Los hallazgos de Nieuwenhuijsen (2006) corroboran los resultados de los estudios anteriores, ya que en sus datos los porcentajes de anteposición muestran un incremento a partir del siglo XIII hasta llegar al siglo XVI de 12% a 54% y 62%; pero concluye que luego este porcentaje baja a 8%, “lo que desemboca en la posposición absoluta con las formas no finitas en el español moderno” (p. 1385). Los trabajos de Berta (1999), y de Luna Traill y Parodi (1974) concuerdan con los de Keniston y de Nieuwenhuijsen en cuanto a las preferencias y tendencias que se observan en la posición del clítico durante los siglos que estudian. Las construcciones con gerundio presentan una situación interesante debido a que ilustran la verificación del proceso inverso; Torres Cacoullós (1999) conduce un estudio diacrónico de estructuras del tipo *-ndo* y demuestra que hay un aumento en la frecuencia de la anteposición de clíticos. Sus conclusiones al respecto se basan en la comparación de textos del español antiguo con variantes del español moderno. Gutiérrez (2008b) estudia una serie de documentos escritos en México y en los Estados Unidos entre los siglos XVI-XVIII para cubrir el período que no fue

estudiado en los trabajos mencionados más arriba; sus resultados presentan una creciente preferencia hacia la posición posverbal a lo largo de los siglos.

En la modalidad oral del español actual, la tendencia en la posición de los clíticos ha adoptado una fuerza creciente a favor de la posición preverbal; pero éste es un proceso que parece estar lejos de resolverse. Este hecho permite estudiarlo en la sincronía, y determinar los factores que favorecen la imposición del cambio y los factores que permiten resistir esta fuerte tendencia. Gutiérrez (2008a), Gutiérrez y Silva-Corvalán (1993), Silva-Corvalán y Gutiérrez (1995), Myhill (1989) han estudiado aspectos del fenómeno; entre las conclusiones de los distintos trabajos aparece cierta sensibilidad a la situación de contacto lingüístico, a los distintos tipos de discurso, al tipo de verbo que funciona como auxiliar y a la naturaleza del verbo que aporta el significado léxico. Troya Déniz y Ma. Pérez Martín (2011) encuentran 83% de subida de clíticos en datos provenientes de 24 grabaciones con estudiantes universitarios en Las Palmas de Gran Canaria. En sus resultados se nota claramente además una diferencia entre las perífrasis de gerundio, que presentan 11 puntos porcentuales más a favor de la subida que las de infinitivo: 91/80). Las autoras comparan sus resultados con los obtenidos en distintas investigaciones con datos del español de Venezuela y concluyen que los resultados de éstas son bastante parecidos a los de Las Palmas, tanto en los porcentajes elevados de subida como en las diferencias entre las perífrasis con infinitivo y gerundio. Aijón Oliva (2010) utiliza los datos del Corpus de Lenguaje de los Medios de Comunicación de Salamanca (MEDIASA), que reúne muestras de la prensa escrita y la radio local, y explora la posición de *se* en construcciones que permiten la subida; en 603 contextos posibles, la subida se da en un 64.2%. El porcentaje es favorable a la subida de clíticos, pero está por debajo de los resultados obtenidos en investigaciones con datos orales, lo que puede explicarse probablemente por los datos escritos que incorpora la base utilizada en esta investigación. De hecho, la naturaleza oral o escrita de los datos a observar ha demostrado ser un factor importante que se debe considerar al estudiar el fenómeno con datos del español moderno; algunos trabajos de las últimas décadas, como el recién comentado, han entregado evidencia que presenta importantes diferencias en la posición de los clíticos entre el registro oral y el escrito. Otro ejemplo se encuentra en Davies (1995), quien estudia tanto documentos escritos como orales; sus resultados entregan información que indica que la subida de los clíticos en los datos orales que estudia es dos veces mayor a la de los datos escritos.

El problema relacionado con las diferencias semánticas entre ambas posiciones ha sido comentado con mayor o menor detenimiento prácticamente en todos o casi todos los estudios sobre la subida de clíticos. Sin embargo, debido a lo complejo del fenómeno, más aportes se requieren en esta discusión. Como dijimos más arriba, en el presente estudio consideramos que el contexto que se estudia es variable; sin embargo, nos interesa explorar varios factores que podrían incidir en la elección que hacen los hablantes.

## 2. METODOLOGÍA

Los datos utilizados en el análisis están constituidos por noventa conversaciones grabadas en video de alrededor de cuarenta minutos de duración con igual número de hablantes. Las conversaciones guiaban a los hablantes a desempeñarse lingüísticamente en distintos tipos de discurso y fueron conducidas durante la última parte del año 2009 entre estudiantes previamente entrenados y hablantes mexicanoamericanos de la ciudad de Houston. Los hablantes fueron agrupados en tres generaciones de acuerdo a la edad de adquisición del español e inglés, o al número de años de contacto con esta lengua, a los años de escolaridad en español y a la edad en que llegaron a los Estados Unidos, si éste era el caso. De esta manera, los hablantes nacidos en México que emigraron a USA después de los 11 años de edad formaron parte de la primera generación, los hablantes nacidos en Estados Unidos o que emigraron a este país desde México antes de los 6 años de edad se consideraron como de segunda generación, mientras los hijos de hablantes de este último grupo están dentro de la tercera generación. Cada grupo generacional está compuesto por un número similar de mujeres y hombres (9/10, 16/18, 17/20, total: 48 mujeres y 42 hombres). Esta agrupación permitió estudiar el fenómeno utilizando el tiempo aparente. Los ejemplos de habla real citados a lo largo del trabajo contienen la información sobre la procedencia de los datos (H: Houston), la generación (1, 2, 3) y el sexo de los hablantes (hombre: 1, mujer: 2). Un análisis llevado a cabo con datos de la misma comunidad recogidos a principios de la década de los noventa permitió tener resultados de dos momentos separados por casi veinte años; la comparación entre éstos constituye el estudio del fenómeno en tiempo real.

La transcripción de las conversaciones grabadas facilitó más tarde la identificación de las estructuras que constituyen el objeto de este estudio.

Cada una de las formas seleccionadas fue codificada de acuerdo a una serie de variables lingüísticas y sociales; en lo que concierne a estas páginas se discutirá la variable dependiente (posición del clítico) en relación al tipo de frase verbal (infinitiva o de gerundio), al verbo auxiliar de la construcción, a la generación de los hablantes y al tiempo histórico.

Las frases verbales seleccionadas fueron aquellas que contenían un verbo conjugado que funcionaba a modo de auxiliar (algunos con preposición) más una forma en infinitivo o gerundio, y uno o dos clíticos adherido/s a estos últimos o antepuesto/s al verbo conjugado. Para la identificación inicial en el corpus electrónico de formas que podrían ser consideradas en el análisis se utilizó el programa Wordsmith (1996); una revisión posterior de cada contexto aislado por este programa permitió hacer la selección definitiva de las formas a considerar en el estudio. Para el análisis estadístico de los resultados se utilizó SPSS (2010). Como se puede apreciar en los ejemplos (3) - (5), se encontró una variedad amplia de verbos gramaticales que permite la subida de los clíticos.

- (3) ... si **me puedo desarrollar** de una manera de la que **me estoy desarrollando** aquí. (H/1/1)
- (4) ... entre mientras más tiempo va avanzando **se va perdiendo** más el idioma. (H/3/2)
- (5) ... no habla muy bien el inglés, ah, así es que a veces **tengo que hablarle** en de alguna manera ... (H/2/2)

Un total de 1210 clíticos fueron incorporados al estudio; 907 forman parte de frases verbales con infinitivo y 303 con gerundio.

### 3. RESULTADOS

Los resultados generales indican que los hablantes favorecen en gran proporción la subida del clítico en aquellos contextos en que existe la opción de elegir el lugar en que éste puede ubicarse. Casi dos tercios de la estructuras examinadas, 779, presentan un clítico antes del verbo que funciona como auxiliar, mientras 431 de éstas mantienen el clítico junto al verbo léxico, un 64.4% frente a un 35.6% que ilustra la gran preferencia de los hablantes por la posición preverbal. El porcentaje de subida se encuentra 6.7 puntos porcentuales por debajo de lo encontrado hace veinte años (71.1%, 1103/1551); sin embargo, la variedad y cantidad

de factores que pueden estar afectando la selección de los hablantes recomiendan interpretar con cautela esta distancia. Hay diferencias claras según el tipo de estructura que se considere; las perífrasis de gerundio tienden a mantener la subida de clíticos casi en forma categórica, pero las de infinitivo presentan un continuum que va de algunas perífrasis que la favorecen, poniéndolo junto al verbo que cumple una función gramatical, hasta otras que privilegian la posición de este elemento junto al verbo léxico al final de la perífrasis. La variedad de estructuras en que se observa la subida de clíticos y lo variado del comportamiento de los hablantes en los distintos tipos de frases que la permiten, obligan a conducir un análisis que explore más de cerca la naturaleza de cada uno de los contextos en que se presenta el fenómeno.

### 3.1. Gerundio vs infinitivo

Hay un claro predominio de la subida del clítico en aquellos contextos en que la perífrasis verbal contiene un gerundio como verbo léxico y se podría decir que este fenómeno es prácticamente categórico. La Tabla 1 presenta un panorama que no deja dudas al respecto.

Generación	% de subida	Casos	Total
Primera	94	78	83
Segunda	90.7	98	108
Tercera	96.4	108	112
Total	93.7	284	303

$$X^2 = 3.038, p .218$$

**Tabla 1.** Subida y generación en frases de gerundio

Casi un 94% de la totalidad de contextos con gerundio examinados (303) lleva el clítico en la posición preverbal. Esto deja ver con claridad que la posición posverbal casi no es usada por los hablantes para posicionar este elemento; este comportamiento de los hablantes es uniforme a lo largo de las generaciones, pues en todas ellas la subida está casi por sobre los noventa puntos porcentuales (94, 90.7 y 96.4). La escasa diferencia que se observa entre los grupos es confirmada por la prueba estadística, por lo que la variable no constituye un factor determinante en la elección de la posición del clítico; la importancia de estos resultados radica en la magnitud que alcanza la subida en las tres generaciones. Esta situación muestra poca diferencia con

la que se encontró veinte años atrás; entonces había 88.6% (241/272) y las generaciones mostraban porcentajes similares: 86.9, 88.3 y 87.7. Aunque pequeña, esta diferencia a favor de la subida en los datos más recientes indica, sin embargo, que cualquier variación que haya en el contexto ocurrirá dentro del dominio casi absoluto de la subida del clítico.

En los contextos con verbo léxico en infinitivo, la situación es bastante diferente a la discutida en la sección anterior. El porcentaje de subida de clíticos es más bajo, llega a poco más de la mitad de las estructuras que aceptan el fenómeno estudiado. Mientras 54.6% presenta clíticos antes del verbo gramatical, el 45.4% restante mantiene el clítico junto al verbo léxico. Las estructuras con infinitivo parecen ser menos estables también en cuanto a este fenómeno, ya que hay una clara diferencia con lo que ha ocurrido en la comunidad en el lapso de veinte años que separa los dos grupos de datos examinados. Se encontró un 68.4% (969/1417) de subida de clíticos en los datos de principios de los noventa; de acuerdo a esto, el fenómeno ha disminuido en 13.8 puntos porcentuales. No fue posible encontrar una tendencia en los datos anteriores, ya que la diferencia de subida de clíticos entre generaciones era mínima: 66.1, 67.8, y 69.7. La situación actual se puede apreciar en la Tabla 2.

Generación	% de subida	Casos	Total
Primera	50	112	224
Segunda	54.5	180	330
Tercera	57.5	203	353
Total	54.6	495	907

$$X^2 = 3.116, p .210$$

**Tabla 2.** Subida y generación en frases de infinitivo

Los porcentajes presentados en la Tabla 2 indican una leve tendencia hacia el aumento de la subida de clíticos, ya que mientras la primera generación de hablantes muestra una mitad de estructuras a cada lado de las perífrasis, la segunda aumenta a un 54.5% los casos de subida de clíticos y la tercera aumenta un poco más a un 57.5%. La tendencia moderada podría ser indicadora de un cambio ya que las nuevas generaciones favorecen esta posición. Sin embargo, la información de los datos de hace veinte años hace suponer que el contexto no es estable a este respecto. En realidad, lo que hemos visto en las últimas décadas es un retroceso del fenómeno; ni siquiera

las frecuencias superiores en más de 13 puntos porcentuales se mantuvieron con estabilidad en el último tiempo. Aun cuando se tienen en cuenta resultados de dos períodos en tiempo real y dos momentos con datos en tiempo aparente, los resultados indican que deben ser tomados con mucha cautela. La situación recién descrita no permite predecir lo que pasará en las próximas décadas con este fenómeno a nivel general, ya que es posible que estemos en presencia de una variación estable; esto es corroborado por la prueba estadística empleada, que no manifiesta diferencias significativas entre los grupos generacionales. Además, investigaciones anteriores han presentado evidencia que indica que en algunas épocas los hablantes han favorecido una posición mientras en otras han mostrado preferencia por la otra. Se hace necesario, por lo tanto, examinar lo que ocurre con las perífrasis específicas y recurrir a datos diacrónicos de más largo plazo; un examen como éste puede ofrecer mayor información sobre este complejo fenómeno.

### 3.2. Auxiliar

Las diferencias que se observan en la subida del clítico según el verbo que ofrece la información gramatical en la perífrasis indican que este elemento es una variable que incide en gran medida en la elección de los hablantes. Este hecho se confirma con el resultado de la prueba estadística empleada.

Como se puede apreciar en la Tabla 3, casi todos los auxiliares que aparecieron en las perífrasis con gerundio (*andar, quedar, venir, ir y estar*) no sólo aceptan subida del clítico, sino que la favorecen categóricamente con frecuencias por encima de los 95 puntos porcentuales. En este contexto la situación es bastante clara y podría estar ligada al hecho de la sílaba extra que introducen las formas de gerundio, aunque se requiere contar con más información acerca de este asunto antes de poder afirmarlo con mayor seguridad. La única forma que parece atraer menos al clítico a un lugar inmediatamente anterior al suyo es *seguir*. Es posible que algún rasgo del significado léxico que mantiene, especialmente la idea de recurrencia en transcurso, lleve a los hablantes a identificarla como el elemento gramatical en que los clíticos se acomodan de manera más apropiada, pero este proceso estaría aún en progreso. En cualquier caso habría que tener en cuenta dos hechos que surgen de los resultados obtenidos. En primer lugar, en el contexto particular de las perífrasis de gerundio con *seguir* como auxiliar, la preferencia por la subida es marcada, de dos a uno prácticamente (63.6%

v/s 36.4%). En segundo lugar, el número de casos encontrados es bajo; hubo sólo 11 instancias en que los hablantes produjeron este tipo de construcción. Cualquier explicación que se ofrezca es, por lo tanto, exploratoria y necesita apoyarse en un mayor número de datos.

Posición →	% de subida	Casos	Total	% G: 1/2/3
<b>Auxiliar Ger ↓</b>				
<i>andar</i>	100	7	7	
<i>quedar</i>	100	5	5	
<i>venir</i>	100	4	4	
<i>ir</i>	96.4	27	28	100/50/100
<i>estar</i>	95.5	234	245	96/94/97
<i>seguir</i>	63.6	7	11	
<b>Auxiliar Inf ↓</b>				
<i>poder</i>	82.7	301	364	79/85/83
<i>querer</i>	59.4	101	170	65/53/63
<i>ir a</i>	51.4	54	105	41/54/51
<i>saber</i>	50	16	32	50/58/42
<i>deber/de</i>	32.1	9	28	17/20/50
<i>pensar</i>	18.2	2	11	
<i>venir</i>	7.1	1	14	
<i>empezar/a</i>	4.8	1	21	
<i>tratar de</i>	3	1	33	
<i>tener que</i>	2.2	2	91	

$\chi^2 = 300.911$ ,  $p .000$

**Tabla 3.** Porcentajes de subida de clíticos y auxiliar (frecuencias de 4 o más)

Los resultados obtenidos también dejaron en evidencia otras tres formas de frecuencia escasa en los datos examinados (*andar*, *quedar* y *venir*), pero la totalidad de contextos en que aparecieron como verbos auxiliares presentó clíticos antepuestos a ellas. Las dos formas más



frecuentes, *ir* y *estar*, manifestaron porcentajes prácticamente categóricos a favor de la subida del clítico. La diferencia en el caso de *ir* es clara, ya que de 28 veces en que funcionó como componente gramatical de una perífrasis de gerundio en 27 el clítico fue producido antepuesto a éste. Pero el caso más ilustrativo es, sin duda, *estar*. Esta forma parece ser la más gramaticalizada de todas, ya que sus rasgos léxicos sólo aparecen de manera esporádica junto a algunas construcciones. En el contexto que se estudia, *estar* aparece además con frecuencias muy elevadas debido al significado progresivo que con frecuencia usan a diario los hablantes de español. Las 245 construcciones progresivas parecen dejar claro que no debe haber duda sobre el lugar que ocupa el clítico en ellas; podrían indicar además que los hablantes favorecen una función más gramatical del clítico al alejarlo del verbo que expresa la mayor carga léxica de la perífrasis. (6) - (9) son ejemplos ilustrativos de subida de clíticos con distintas perífrasis de gerundio.

- (6) ... [no tiene] todos los papeles derechos como es aquí, pero aquí la policía **nos viene instruyendo** para, para que sea, sigamos las reglas ... (H/1/1)
- (7) ... un hombre que tiene mucho dinero ... y su hermano **se anda divorciando** su esposa ... (H/2/1)
- (8) ... la experiencia más bonita que he tenido aquí, **la sigo teniendo** seguido, es tener toda mi familia junto conmigo ... (H/2/2)
- (9) ... I mean porque allí podía ver el carro y yo estaba en el, ya **me iban metiendo** en el helicóptero y I mean eso era, taba llorando y llorando (H/3/2)

Podría ser que en estos contextos la posición del clítico antes del verbo auxiliar realce el significado gramatical, más abstracto, de la forma que funciona como auxiliar. En (6) *viene* no tiene la intención de indicar un movimiento físico que ejecuta la policía, sino que nos habla más bien de una costumbre que se ha hecho frecuente en ese ambiente social. En (7) *anda* no se refiere a un momento en que el hermano del sujeto del discurso del hablante haya ido a divorciarse de su esposa, sino a una etapa temporal más amplia que tiene que ver más bien con el aspecto verbal. Fenómenos similares ocurren en (8) y (9), pues los clíticos antepuestos forman parte de perífrasis progresivas en que este aspecto pareciera tener mayor relieve con la presencia del clítico al principio de éstas.

Por otro lado (10) - (16) presentan contextos de perífrasis en que el significado léxico está dado a través de una forma en infinitivo. Los auxiliares de (12), (13) y (15) son los únicos, sin embargo, que ofrecen porcentajes de importancia a favor de la subida de clíticos. Parece que en estas formas, el significado léxico se ha debilitado más que en las restantes; esto se ve claramente en el caso de *ir + a*, ya que ambas frases verbales en (13) constituyen perífrasis de futuro y no de movimiento, significado que poco a poco va dejando de lado esta forma lingüística. Aunque es menos claro, es posible que esté ocurriendo un fenómeno equivalente en estructuras en que aparecen formas del tipo *saben* y *debió*.

- (10) ... meterme yo a educarme más uhm y y sí **lo pienso hacer** porque ... (H/3/2)
- (11) ... pero no me hacer sentir así como que ya ... sí, sí, sí, **me va venir** la buena suerte. (H/3/2)
- (12) ... si no **lo saben defender** se va muy lejos y pasa la mesa, ... (H/3/1)
- (13) ... o en cualquier lado de Estados Unidos y pues **te va [a] pagar** los viajes de tus niños, **te voy a pagar** que te muevas de casa ... (H/1/1)
- (14) ... y también el agua **se empezó a meter** al restaurante entonces nos subimos a las mesas ... (H/1/2)
- (15) ... fue una guerra que no **se debió perder** pero se perdieron muchas almas jóvenes de mis escuela ... (H/2/2)
- (16) ... no supe qué me estaba diciendo y **le tuve que preguntar**. (H/2/1)

La subida del clítico en frases verbales con infinitivo funciona de manera muy diferente a como ocurre en las frases con gerundio. Esta situación queda en evidencia al comparar las distintas frases que aceptan la subida en los datos examinados; como se aprecia en la Tabla 3, sólo hay un par de éstas que favorece de manera clara este fenómeno: *poder* y *querer*. Sin duda, *poder*, en su papel de auxiliar presenta un terreno fértil que permite a los hablantes desprender el clítico del verbo principal de la frase. En 364 instancias de frases con *poder*, los hablantes eligieron esta posición 301 veces, lo que equivale al 82.7%; sólo hay un 17.3% de contextos (63 casos) en que los hablantes han mantenido el clítico junto al verbo que aporta el significado léxico. Esta situación se mantiene, con pequeñas diferencias, en las tres generaciones, por lo que no cabe duda que se va a

mantener más cerca de las perífrasis de gerundio que de otras de infinitivo. Las perífrasis con *querer* como auxiliar favorecen en gran medida la subida del clítico, ya que éstas alcanzan casi un 60% de las 170 estructuras que permitían la subida. Sin embargo, éstas se encuentran por debajo del nivel de selección por la subida que hacen los hablantes con *poder*. En el caso de *ir a* y *saber* se observan porcentajes elevados de subida, pero las frases sin subida se mantienen en un nivel similar; los hablantes no se inclinan por ninguna de las dos situaciones. Con un porcentaje menor, se aprecia que *deber/de* tiene un tercio de subida de clítico. La Tabla 3 deja en claro también que hay una serie de frases de este tipo que no favorece la subida de clíticos; la preferencia de los hablantes en estos casos ubica los pronombres mayoritariamente adheridos a la forma verbal en infinitivo que aporta la información léxica a la perífrasis.

Resulta difícil establecer diferencias claras que permitan explicar las preferencias de los hablantes por favorecer la subida del clítico en ciertas perífrasis y no en otras. Con claridad se puede ver, sin embargo, que las perífrasis progresivas favorecen esta elección por parte de los hablantes. Un efecto similar producen las perífrasis con verbos auxiliares como *poder*, *querer*, y en menor medida *ir a*, *saber* y *deber/de*. No se puede ser categórico al momento de sacar conclusiones al respecto, pero algunas de estas formas han experimentado un proceso de gramaticalización que los ha especializado en su función de auxiliares. Es posible que el pronombre clítico que se refiere a los objetos sea visto por los hablantes como una categoría más gramatical que léxica en estos casos; con esto, el lugar propio es más bien junto al elemento que ofrece las categorías gramaticales y no léxicas. Este hecho podría explicar la fuerte preferencia preverbal de estas formas en perífrasis con verbos modales como *poder* y *querer*, como las que aparecen en (17) y (18).

- (17) ... también tener una protección que **les puede ayudar**  
 pa que ellos **pueden educarse** y sacar esa su familia  
 adelante. (H/3/2)
- (18) ... no **les quiero poner** la culpa a todos pero no, I mean,  
 nunca tuve ambición de la escuela y ... (H/2/1)

A nivel diacrónico se pueden notar algunos cambios, según se observa en la Tabla 4 que presenta los resultados de investigaciones llevadas a cabo en tres momentos diferentes en los últimos treinta años.

Época →	Houston 10s	Houston 90s	Los Ángeles 80s	Mich 80s
<b>Auxiliar ↓</b>				
<i>estar</i>	96	88	91	93
<i>poder</i>	83	77	74	81
<i>querer</i>	59	52	45	77
<i>ir a + inf.</i>	51	86	93	83
<i>deber</i>	32	57	38	100 (1/1)
<i>empezar/com a</i>	5	65	71	100 (2/2)
<i>tener que</i>	2	55	63	29

Hou 90s:  $X^2 = 283.030$ , p .000

Mich:  $X^2 = 33.631$ , p .009

**Tabla 4.** Porcentajes de subida de clíticos y auxiliar en Houston, Los Ángeles y Michoacán

Los datos de Houston de las distintas épocas permiten ver alzas en la subida del clítico en las construcciones de infinitivo que más favorecen este proceso. Las formas son las mismas, tanto en los datos de los noventa como en los del 2010, *estar*, *poder* y *querer*, y no sólo continúan atrayendo las preferencias de los hablantes con respecto a la subida, sino que presentan avances de importancia. Ocho puntos porcentuales más presentan las perífrasis de gerundio con *estar* (88 a 96), seis puntos es el aumento en el caso de las perífrasis de *poder* más infinitivo (77 a 83) y siete en el caso de *querer*, aunque en el caso de estas últimas el apoyo es de poco más de la mitad de las construcciones (52 a 59). *Estar* y *poder* también presentaban frecuencias muy altas en los ochenta, tanto en Los Ángeles como en la comunidad monolingüe de Michoacán, pero el aumento de la subida en los datos más recientes de Houston está por sobre las frecuencias de esa época. *Querer* ha mostrado un progreso sostenido a lo largo de las décadas, ya que en ambos períodos anteriores las frecuencias en la subida del clítico ha sido menor (45 a 52 a 59); sin embargo, la variedad monolingüe presenta una frecuencia aun mayor (77%).

Los resultados presentados a lo largo de los años permiten apreciar una disminución de la subida del clítico en varias perífrasis de infinitivo. Esta situación se observa en los casos en que el verbo auxiliar es *ir a*, *deber*, *empezar/comenzar* y *tener que*. Tanto en Houston en los datos de los años

noventa, como en Los Ángeles una década antes, estas formas mostraban frecuencias considerables que indican que los hablantes favorecían la subida de clíticos. Llama la atención *ir + a*, ya que presenta una frecuencia alta en las entrevistas de los hablantes, lo que le da mayor confiabilidad a los resultados relacionados con esta forma, que manifiesta una fuerte caída a lo largo de los años (86% y 93%) hasta llegar a un 51% con construcciones con subida de clíticos, 54 de 105 construcciones posibles. La proporción es aún de bastante importancia, pero ha llegado a la mitad de todas las construcciones posibles; habrá que esperar a futuro para ver qué ocurre en esta estructura que se ha especializado, y es altamente preferida por los hablantes, en la expresión de futuridad con respecto a un punto anterior. Las perífrasis con *empezar/comenzar + prep y con tener que* tenían en las décadas pasadas frecuencias bastante favorables a la subida de clíticos; sin embargo, los datos más actuales revelan una caída significativa, con sólo un 5 y un 2 por ciento, 1 en 21 contextos posibles en el primer caso y 2 en 91 en el segundo. Una forma que mantiene cierta proporción con subida de clíticos es *deber*, aunque también ha experimentado una baja con respecto a las décadas anteriores.

#### 4. CONCLUSIÓN

Los datos examinados a lo largo del tiempo real confirman que el fenómeno de subida de clíticos no puede analizarse de manera global, sino que se debe considerar cada perífrasis o tipo de perífrasis por separado una vez que el uso ha indicado, en los datos disponibles, cuáles aceptan la subida de clíticos. En este caso, queda claro que hay un grupo que no sólo se ha mantenido por décadas favoreciendo de manera casi categórica la subida de clíticos, incluso muestra un aumento en esa misma dirección; las perífrasis progresivas con *estar + gerundio y poder + infinitivo* han sido identificadas de esta manera. Hay otro grupo que se ha mantenido con un uso intermedio y muestra ya una tendencia a aumentar ya una tendencia a disminuir; este hecho hace que sea muy difícil predecir si las perífrasis con *querer* y con *ir a + infinitivo* seguirán una u otra de las tendencias indicadas con respecto a la subida de clíticos. Hay algunas perífrasis también que han dejado de ser favorables a la subida de clíticos, ya que muestran una disminución clara con respecto a las décadas anteriores; es el caso de aquellas que se construyen con los auxiliares *empezar/comenzar + prep y con tener que*.

Los resultados obtenidos en tiempo aparente a partir de tres grupos de hablantes de acuerdo a la variable generación, se ajustan a los recién comentados. Las dos perífrasis que más favorecen la subida de clíticos, *estar* + gerundio y *poder* + infinitivo, se mantienen, con pequeñas variaciones, con frecuencias muy altas de subida de clíticos a lo largo de las tres generaciones. Las frecuencias que presentan (96, 94 y 97 la primera y 79, 85 y 83 la segunda) casi no permiten seguir aumentando; en realidad, un fenómeno tan variable como el que se examina probablemente nunca va a manifestar un comportamiento totalmente categórico. La evidencia en el tiempo real y aparente permite ubicarlas en un nivel similar al actual en cuanto al fenómeno examinado. Las perífrasis con *querer* y con *ir* + *a* + infinitivo en tiempo aparente presentan resultados más variables, pero de gran interés. La primera mostraba una tendencia al aumento de subida en tiempo real y el tiempo aparente indica también lo mismo, aunque presenta una disminución momentánea en el caso de la segunda generación, la tercera generación, sin embargo, retoma con fuerza la tendencia a la subida con porcentajes de 65, 53 y 63 respectivamente. La segunda forma mostró una clara disminución en tiempo real; en el tiempo aparente, sin embargo, parece consolidar las preferencias de manera balanceada entre subida y no subida, ya que mientras la primera generación presenta un 41% de subida de clíticos, este fenómeno ocurre en la segunda con una frecuencia de 54%; la tercera generación, con un 51%, representa el promedio de toda la comunidad de Houston en el momento presente. Las perífrasis con *deber* muestran una tendencia al alza de la subida de clíticos, con porcentajes de 17, 20 y 50 a lo largo de las generaciones, pero tanto el bajo número de casos con subida encontrados como los porcentajes que estas perífrasis alcanzan en tiempo real no permiten sacar mayores conclusiones. Lo mismo ocurre con perífrasis con *empezar/comenzar* + prep y con *tener que*, en oposición a los resultados en tiempo real, presentan frecuencias muy bajas en los datos de los últimos años en la comunidad de Houston (1/21 en el primer caso y 2/91 en el segundo).

Los resultados obtenidos presentan evidencias claras de que el problema de la posición del clítico en un grupo de perífrasis verbales que permiten la posición variable de éste sólo se encuentra más o menos resuelto en algunas de éstas. La mayoría de ellas, sin embargo, ofrece al hablante ambos espacios; pero los datos reunidos hasta el momento sólo permiten afirmar que el problema, si realmente lo es, está lejos de resolverse.



## REFERENCES

- Aijón Oliva, Miguel (2010): “La posición de *se* en las estructuras pluriverbales: variación y significado”, *Boletín de Filología* XLV-22, pp. 11-29.  
<http://dx.doi.org/10.4067/S0718-93032010000200001>
- Berta, Tibor (1999): “La Celestina: Desde el castellano medieval hacia el español clásico”, *Acta Hispánica* IV, pp. 69-80.
- Davies, Mark (1995): “Analizing syntactic variation with computer-based corpora: the case of Modern Spanish clitic climbing”, *Hispania* 78, pp. 370–80.  
<http://dx.doi.org/10.2307/345438>
- Gutiérrez, Manuel J. (2010): “La media subida de clíticos en el español de Houston”, *Nueva Revista de Filología Hispánica* 58-2, pp. 645-662.
- (2008a): “Restringiendo la subida de clíticos: reflexividad, modalidad verbal y contacto lingüístico en el español de Houston”, *Hispanic Research Journal* 9-4, pp. 299-13.
  - (2008b): “La subida de clíticos durante los siglos XVI-XVIII”, en Concepción Company y José G. Moreno de Alba (eds.): *Actas del VII Congreso Internacional de Historia de la Lengua Española*. Madrid: Arco/Libros, S.L., pp. 641-655.
  - (2008c): “Subida de clíticos, tipo de frase verbal, reflexividad, generación y género en el español de los Estados Unidos en los siglos XIX y XX”, *Revista Iberoamericana de Lingüística* 2-3, pp.29-48.
  - y Carmen Silva-Corvalán (1993): “Clíticos del español en una situación de contacto”, *Revista Española de Lingüística* 23, pp. 207–220.



- IBM (2010): *SPSS = Instrucciones para el uso del programa SPSS*. Internet: [ucm.es/info.manual.spss](http://ucm.es/info.manual.spss).
- Keniston, Hayward (1937): *The syntax of the Castilian prose: The sixteen century*. Chicago: The University of Chicago Press.
- Lexical Analysis Software Ltd. y Oxford University Press (1996): *WordSmith 5.0*. Oxford: Oxford University Press.
- Luna Traill, Elizabeth y Claudia Parodi (1974): "Sintaxis de los pronombres átonos en construcciones de infinitivo durante el siglo XVI", *Anuario de Letras* 12, 196-204.
- Myhill, John (1989): "Variation in Spanish Clitic Climbing", en Thomas J. Walsh (ed.): *Synchronic and diachronic approaches to linguistic variation and change*. Washington D. C.: Georgetown University Press, pp. 227-250.
- Nieuwenhuijsen, Dorien (2006): "Cambios en la colocación de los pronombres átonos", en Concepción Company (ed.): *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*. México, D.F.: UNAM/Fondo de Cultura Económica, pp. 1339-1404.
- Silva-Corvalán, Carmen y Manuel J. Gutiérrez (1995): "On Transfer and Simplification: Verbal Clitics in Mexican-American Spanish", en Peggy Hashemipour, Ricardo Maldonado y Margaret Van Naerssen (eds.): *Studies in Language Learning and Spanish linguistics in honor of Tracy D. Terrell*. San Francisco: McGraw-Hill, pp. 302-12.
- Suñer, Margarita (1980): "Clitic Promotion in Spanish Revisited", en Frank H. Nuessel Jr. (ed.), *Contemporary Studies in Romance Languages*. Bloomington: Indiana University Linguistics Club, pp. 300-30.
- Torres Cacoullós, Rena (1999): "Construction Frequency and Reductive Change: Diachronic and Register Variation in Spanish Clitic Climbing", *Language Variation and Change* 11, pp. 143-170.  
<http://dx.doi.org/10.1017/S095439459911202X>
- Troya Déniz, Magnolia y Ana M. Pérez Martín (2011): «Distribución de clíticos con perífrasis verbales en hablantes universitarios de Las Palmas de Gran Canaria», *Lingüística* 26, pp. 9-25.



## **Lateralización de /r/ implosiva: la conciencia fonológica y sus manifestaciones en el español puertorriqueño oral y escrito**

| Antonio Medina-Rivera  
| *Cleveland State University*

**Abstract:** This study examines the process of lateralization of syllable final /r/ in Puerto Rican Spanish. In addition to analyzing the process in oral speech, this study takes into consideration written registers as well. The results show that the process of lateralization goes beyond oral registers among Puerto Rican Spanish. However, the manifestation of the lateral variant does not only occur as a result of neutralization or confusion between ‘r’ and ‘l.’ There seems to be a ludic function and evident intentionality in the use of lateral variant, that in some ways shows a mark of identity among Puerto Rican speakers. The study incorporates written samples from several individuals as well as examples from Facebook.

**Keywords:** Lateralization, Puerto Rican Spanish, written registers, Spanish in the Internet, stylistic/register variation.



## **1. INTRODUCCIÓN**

Los trabajos fonológicos sobre el español de Puerto Rico ponen especial atención a los procesos de lateralización de /r/ implosiva, velarización de la vibrante múltiple y la aspiración o elisión de /s/ implosiva. Por razones obvias, los investigadores toman en consideración los registros hablados para examinar la frecuencia de dichos fenómenos entre hablantes puertorriqueños, y los tres fenómenos mencionados se han convertido en los elementos principales para caracterizar el español de la Isla. Este trabajo pretende hacer un recorrido a estudios fonológicos de carácter cuantitativo del pasado, y a la vez extender el tema a las maneras en la que el fenómeno de lateralización se puede expresar en registros escritos, ya sea como una función lúdica o como una manifestación variable en la escritura de un hablante de español puertorriqueño.

Para analizar el habla escrita se contará con trabajos escritos a mano de hablantes puertorriqueños y con ejemplos tomados de Facebook. Los trabajos escritos a mano y los mensajes espontáneos de Facebook no cuentan con un mecanismo de autocorrección, por lo tanto pueden ser útiles para examinar la lengua de la manera más natural posible. No se pretenderá cuantificar los ejemplos escritos de la misma manera en la que se cuantifica una entrevista sociolingüística, debido a las limitaciones que existen para recolectar este tipo de ejemplos. Los datos que nos proveen los estudios fonológicos de varios investigadores y los datos que se ofrecerán de los medios escritos servirán para tener una idea más completa de una variable fonológica que forma parte integral del repertorio lingüístico puertorriqueño.

## **2. CONCIENCIA, PERCEPCIÓN Y ACTITUDES LINGÜÍSTICAS**

El concepto de conciencia fonológica sería esencial para el análisis de este estudio ya que nos ayuda a entender el conocimiento y apreciación que los hablantes de una variedad específica tienen sobre su dialecto local. Quisiéramos creer que cada hablante es realmente consciente de las variaciones que existen en su repertorio fonético - fonológico. Por ejemplo, muchos hablantes de español están conscientes sobre la existencia o ausencia de la distinción entre /s/ y /θ/ en su dialecto, o sea si pueden distinguir entre palabras como 'casa' y 'caza' en el habla oral, aunque no conozcan con claridad las reglas que dictan su uso (e.g. saber que la 'z' se pronuncia como /θ/, pero

desconocer que la ‘c’ frente a ‘e’ o ‘i’ se pronuncia de igual manera). Siempre se encontrarán hablantes que no están completamente seguros al respecto o que dicen que hacen una cosa, pero en su actuación lingüística hacen otra.

La gran mayoría de los trabajos sobre conciencia fonológica ponen énfasis en el proceso de aprendizaje de la lectura en niños en sus primeros años de escuela primaria, ya que esta conciencia que tiene el niño va a tener un efecto en su producción oral. El nivel de conciencia lingüística y las percepciones que los individuos (no especializados en lingüística) tienen sobre su lengua o sobre su variedad de habla se acomoda muy bien dentro del marco teórico de la lingüística perceptual. Irania Malaver (2002: 181) define la conciencia lingüística como:

‘el saber’ que sobre los hechos de la lengua tienen los miembros de una comunidad; este saber les brinda no sólo la oportunidad de identificar los usos lingüísticos que poseen mayor reconocimiento cultural dentro de su comunidad sino que también refleja el grado de instrucción o nivel cultural que los miembros de las comunidades poseen.

Dennis Preston (1999) establece un modelo sobre los usos del habla que incorpora (a) lo que la gente dice, (b) lo que la gente dice sobre eso, y (c) la manera en la que la gente reacciona a lo que dice (Preston 1999: xxiii). El origen de este modelo tiene sus bases en el estudio de Labov de 1966 en el que les pide a los participantes que evalúen el nivel social de un hablante específico de acuerdo a su manera de articular un elemento fonológico. En este sentido en la lingüística perceptual se entiende que el hablante tiene conocimiento de la lengua, es capaz de hacer distinciones dialectales y está consciente de algunas diferencias que se dan en el habla. Preston clasifica este conocimiento como parte del bagaje cultural o folclórico que caracteriza a un grupo dialectal específico y que este conocimiento folclórico puede ser de gran utilidad para dialectólogos, variacionistas y otros sociolingüistas.

Yoshio Mase (1999:72) estudia el nivel de conciencia dialectal en Japón para determinar hasta qué punto los hablantes están conscientes de las diferencias dialectales y cuáles son las fronteras que determinan en dónde empieza un dialecto y comienza otro. Este tipo de estudio, como otros que aparecen en el *Handbook of Perceptual Dialectology* editado por Preston (1999) examina no los resultados del sociolingüista que cuantifica ciertos elementos lingüísticos que se dan en una región u otra, sino que se basan en las percepciones y conocimientos de los residentes de un área específica.

Este tipo de estudio puede servir como punto de partida para otros trabajos de carácter cuantitativo.

Los estudios sobre actitudes lingüísticas contribuyen también a nuestro entendimiento sobre la conciencia lingüística de los hablantes de un dialecto en particular. Este tipo de estudios incorpora por un lado, el conocimiento y actitud que los hablantes tienen sobre aspectos específicos del lenguaje, y además nos dan una idea sobre la autoestima lingüística de los individuos de un dialecto (elementos a y b del modelo de Preston). Los estudios sobre actitudes lingüísticas pueden examinar las percepciones que tienen los hablantes sobre su propio dialecto, pero también las percepciones que tienen sobre otros dialectos (elemento c en Preston). Con respecto a las actitudes y las variedades de una lengua, Niño-Murcia señala que: “Among the national varieties, each group of speakers, however, has a mental hierarchy of dialects. Some are considered “better,” “purer,” or more or less “standard” than others, but loyalty to the norms of one’s own group is a very powerful source of solidarity and is strongly encouraged” (2011: 731). En contraste con la definición de conciencia lingüística, Malaver (2002:182) señala que la actitud lingüística “se define como la valoración sociocultural que los miembros de una comunidad poseen hacia su propia variedad de lengua, hacia otras variedades dialectales de la misma lengua, sus sociolectos así como hacia lenguas distintas”.

Tanto los estudios sobre percepción como los de actitudes lingüísticas sirven para tomar en consideración las valoraciones que tienen los individuos sobre su dialecto o sobre otros dialectos o lenguas. En este estudio se pondrá atención especial al español de Puerto Rico, especialmente a un fenómeno fonológico que caracteriza el español de la Isla. Tanto la conciencia como la actitud lingüística operan en la manera en la que cada individuo se manifiesta a nivel oral o escrito.

### **3. EL ESPAÑOL DE PUERTO RICO**

#### **3.1. Conciencia, percepción y actitudes del español de Puerto Rico**

Valentín-Márquez al referirse al fenómeno de lateralización de /r/ implosiva al igual que al de velarización de /r/ indica, tomando en consideración ambos fenómenos, que “high levels of stigmatization are also present in most of the judgements that Puerto Rican informants who participated in language

attitude tests expressed about these phonological features” (2007:73). Para llegar a esta conclusión Valentín-Márquez toma en consideración varios estudios que analizan las actitudes de hablantes hacia ambos fenómenos. Emmanuelli (1986) indica en su estudio con hablantes del sur de la Isla, que el 59.7% de los encuestados muestra una actitud negativa hacia el fenómeno de lateralización, mientras que Medina-Rivera (1997), en su estudio con hablantes de Caguas, Puerto Rico muestra un porcentaje aún más elevado con 71%.

El fenómeno de lateralización, aun cuando no es exclusivo del español de Puerto Rico, se ha convertido en un rasgo que distingue a la comunidad puertorriqueña en y fuera de la Isla. A nivel anecdótico puedo decir que ocasionalmente no falta el hablante de otro dialecto que se acerca a un puertorriqueño y le dice en tono de burla o a modo de juego: “Tú eres de Puelto Lico”. En este caso el ejemplo es estereotípico y va más allá de las posibilidades reales de la lateralización, ya que “Lico” ni representa un caso de vibrante simple ni se encuentra en posición implosiva. Simonet, Rohena-Madrado y Paz concluyen por otro lado en su estudio que “there seems to be evidence of incomplete neutralization of the two phonemes: while /r/ and /l/ are maximally different in, say, syllable-initial position, their acoustic difference is reduced, but not lost, in syllable-final position” (2008: 84). Esto implica que debido a un proceso de neutralización incompleta, a nivel auditivo, los hablantes de un dialecto pueden percibir los sonidos de una manera diferente a como los perciben otros hablantes de otro dialecto.

Además del elemento estereotípico, el cual refleja la manera en la que los hablantes de otras variedades del español reciben el español de Puerto Rico, se encuentra también el elemento lúdico, el cual muestra la manera en la que los mismos puertorriqueños juegan con sus distinciones fonológicas. El aspecto lúdico del lenguaje se menciona en libros de texto sobre lingüística (e.g. Milton Azevedo 2010), pero no es un elemento de estudio que resulte fácil de examinar, de cuantificar o de producir en una entrevista sociolingüística. En el caso particular de este estudio se destacará más adelante la manera en la que hablantes puertorriqueños perciben la lateralización de una manera lúdica o como una manera de reafirmación nacional.

### **3.2. Metodología**

El presente trabajo combina resultados de estudios anteriores que reflejan registros orales del español puertorriqueños, y los mismos se complementan con la recolección de nuevos datos que reflejan el nivel escrito. Lo que se quiere mostrar es que el fenómeno de la lateralización no se

limita a los registros orales. Los estudios de López-Morales (1983), Medina-Rivera (1997) y Valentín-Márquez (2007) sirven para mostrar los resultados que tres investigadores muestran sobre los registros orales en el español de Puerto Rico. El elemento escrito se representa por medio de datos que se recolectaron en un curso de Antropología Cristiana ofrecido en español en la ciudad de Cleveland (2006) en el cual participaron varios puertorriqueños y otros datos que se obtuvieron en Facebook (2010-2013).

Los dos participantes que se escogieron como parte del curso de Antropología Cristiana nacieron y se criaron en la ciudad de San Lorenzo, localizada al centro-este de la Isla, contaban con un dominio de inglés bastante básico y se desempeñaban la mayor parte del día español. En este sentido, a pesar de que se encontraban residiendo en los Estados Unidos, contaban con características similares a los de los otros estudios en los que se analizan los registros orales. Por otro lado, los datos que se obtuvieron por medio de Facebook corresponden en su totalidad a hablantes que residen en Puerto Rico.

La colección de datos escritos tiene sus limitaciones debido a que hay que contar con ejemplos que se hayan escrito a mano o que no cuenten con un proceso elaborado de edición o corrección. Las muestras que se obtuvieron para el curso de Antropología Cristiana eran tareas que los participantes tenían que completar para el curso, mientras que las muestras de Facebook corresponden a interacciones diarias a las cuales el investigador tenía fácil acceso por ser parte de sus contactos personales. En ambos casos hubo un alto nivel de espontaneidad y las muestras son bastantes similares a las que se pueden obtener en un registro oral.

### **3.3. Lateralización de /r/ en posición implosiva en el español de Puerto Rico**

Antes de proceder a los ejemplos de lateralización en registros escritos, es importante poner en evidencia algunos estudios relevantes que cuantifican los niveles de lateralización en la población puertorriqueña. La Tabla 1 representa los niveles de lateralización en tres estudios realizados en las décadas de 1980, 1990, 2000. Cada uno de estos estudios se realizó en zonas diferentes de la Isla: el de López Morales en la zona metropolitana de San Juan, el de Medina-Rivera en la ciudad de Caguas, localizada en el centro de Puerto Rico a unos 20 minutos al sur de San Juan, y el Valentín-Márquez en la ciudad de Cabo Rojo, localizada al suroeste de la Isla.



López Morales San Juan 1983	Medina-Rivera Caguas 1997	Valentín-Márquez Cabo Rojo 2007	Valentín-Márquez Grand Rapids, MI 2007
34,6%	30,8%	42,3%	30,7%

**Tabla 1.** Frecuencias de Lateralización en Varios Estudios del Español de Puerto Rico

Los estudios de López Morales y Valentín-Márquez recogen datos de hablantes de distintas generaciones y distintos niveles socioeconómicos, mientras que los de Medina-Rivera representan frecuencia de uso entre jóvenes adultos con al menos un año de estudios universitarios. Sin embargo, es notable que las frecuencias sobrepasen el 30% en los tres estudios y podría verse como una variable que es característica del español de Puerto Rico, independientemente del nivel educativo, del género o del grupo generacional. La última columna de la Tabla 1 representa las frecuencias de lateralización para una comunidad puertorriqueña en la ciudad de Grand Rapids, Michigan. Como se puede observar, en la comunidad puertorriqueña de Grand Rapids se mantienen el patrón de lateralización de 30% o más de los otros estudios señalados. Valentín-Márquez (2007: 215-220) menciona otros estudios como de los Ramos-Pellicia (2004) y Lamboy (2004) con niveles de lateralización mucho más altos que los de los estudios que se han realizado en la Isla, 61% y 75%, respectivamente, pero cuestiona las maneras de categorizar las variantes o de contar los casos en ambos estudios. Resulta un poco difícil comparar estudios debido a las diferencias metodológicas, como bien se puede deducir de la crítica de Valentín-Márquez, pero en todos estos casos es evidente que la lateralización de la vibrante simple en posición implosiva es frecuente en el español de Puerto Rico, en y fuera de la Isla.

Muchos puertorriqueños que viven en los Estados Unidos adquieren su español vía oral. Como maestro de español en los Estados Unidos puedo notar que mis estudiantes hispanos que se han criado en los Estados Unidos muestran niveles de lateralización hasta en la escritura, cosa que se esperaría ver mucho menos entre los estudiantes puertorriqueños que han estudiado en la Isla y que han recibido instrucción formal en español. Este comentario se queda a nivel anecdótico, ya que no cuento con estudios disponibles que digan lo contrario o que afirmen mis observaciones. También he podido observar que el nivel de conciencia lingüística es diferente, ya que el puertorriqueño que se ha criado en Estados Unidos posee menos conocimiento sobre el

español y sin lugar a dudas percibe sus diferencias dialectales de una manera diferente al que vive en la Isla. Estas percepciones y manifestaciones pueden ser un tema interesante para futuros trabajos sobre el español de Puerto Rico. En términos de registros o estilos, el hablante puertorriqueño que vive en los Estados Unidos carece de la misma cantidad de registros o estilos que el puertorriqueño que vive en la Isla, nuevamente por la falta de instrucción formal en el idioma.

### 3.4. Manifestaciones de lateralización en la escritura

En el 2006 la oficina del College Board en Puerto Rico publicó un estudio sobre las manifestaciones del español escrito entre maestros que habían tomado el examen de certificación que ofrece el Departamento de Educación (*El Nuevo Día de Puerto Rico* 2006:6). El estudio puso de manifiesto que algunos de los maestros que tomaron el examen de certificación poseían destrezas de escrituras equivalentes a las de un estudiante de noveno grado. Algunos de los ejemplos que se citaron en el periódico *El Nuevo Día* son:

- |        |                           |                               |
|--------|---------------------------|-------------------------------|
| (1) a. | Puertorrico               | por Puerto Rico               |
| b.     | sean perdido              | por se han perdido            |
| c.     | con sigo mismo            | por consigo mismo             |
| d.     | estan importante          | por es tan importante         |
| e.     | departamento de educacion | por Departamento de Educación |
| f.     | sicriata                  | por siquiatra o psiquiatra    |
| g.     | por que ria               | por porquería                 |
| h.     | jenes                     | por genes                     |
| i.     | tolbellino                | por torbellino                |

Los ejemplos que se citan en el periódico muestran diferentes problemas ortográficos, aunque también se observan problemas de concordancia, problemas sintácticos, y problemas de secuencia de verbos. De los ejemplos anteriores, me llamó la atención el ejemplo 1.i, en el cual se puede ver la influencia de un aspecto fonológico en la escritura ya que se registra la palabra ‘tolbellino’ en lugar de ‘torbellino’.

Para poder examinar este aspecto fonológico se utilizaron dos fuentes de análisis para este estudio. La primera consiste de una muestra de trabajos escritos a manos por dos hombres puertorriqueños en la ciudad de Cleveland, Ohio, los cuales escribieron varias muestras para un curso de formación religiosa. La segunda muestra consiste de ejemplos tomados de Facebook a lo largo de tres años.

### 3.4.1. Escritura para una clase

Esta primera muestra examina ejemplos escritos de dos hablantes puertorriqueños residentes en Cleveland, Ohio. Ambos participaban de un curso de Antropología Cristiana proporcionado por la diócesis de Cleveland en el 2006 como parte de un programa de certificación para ministros de pastoral. El curso se ofrecía 100% en español y los participantes tenían que trabajar en dos tareas semanales: una reflexión en un diario personal en la cual reaccionaban al tema que se había presentado en la última clase, y una tarea de investigación en la cual se les requería analizar algún tema en específico. El instructor del curso también incluía pruebas cortas y presentaciones orales para los participantes. Los dos participantes que se seleccionaron para el estudio fueron Pablo y Francisco. En el momento en el que se seleccionaron las muestras para el estudio Pablo tenía 39 años y Francisco 41.

Pablo y Francisco son hermanos y completaron la escuela secundaria ('high school') en San Lorenzo, Puerto Rico. Luego de completar la escuela secundaria ambos se vinieron a los Estados Unidos para trabajar en construcción y en jardinería. Ninguno de los dos prosiguió estudios universitarios, y ahora que están en su adultez decidieron recibir formación religiosa con la diócesis de Cleveland. Este tipo de experiencia educativa ha llevado a que Pablo y Francisco reactiven sus destrezas escolares, sobre todo la escritura y la lectura, las cuales habían permanecido un tanto inactivas desde que habían completado la escuela secundaria. Tanto Pablo como Francisco pueden comunicarse en inglés de una manera bastante básica; sin embargo se desenvuelven principalmente en español en la mayoría de sus dominios lingüísticos.

Originalmente se analizaron los usos de /s/ implosiva y de /r/ implosiva en tres tipos de muestras escritas a mano: reflexiones del diario, tareas de investigación y pruebas cortas. Para este estudio solamente se mostraron los ejemplos de /r/ implosiva. La Tabla 2 muestra los resultados de las frecuencias de cambio de 'r' por 'l' en ambos participantes:

	<b>Francisco</b>	<b>Pablo</b>
<b>Reflexiones de Diario</b>	(1/52) 2%	(3/34) 9%
<b>Tareas de Investigación</b>	(0/84) 0%	(1/61) 2%
<b>Pruebas Cortas</b>	(0/22) 0%	(4/10) 40%

**Tabla 2.** Cambio de 'r' por 'l' en muestras escritas

Tal y como se puede ver en la tabla 2, Francisco solamente muestra cambio de ‘r’ por ‘l’ en las reflexiones del diario, mientras que Pablo muestra este tipo de alteración en los tres registros escritos, sobre todo en las pruebas escritas. Algunos de los ejemplos que aparecieron en los textos son:

- (2) a. en sudol
- b. descubril cosas
- c. moril

La gran mayoría de los casos ocurrió en verbos infinitivos como en (2c). No se dieron casos de rotacismo (cambio de “l” por “r”), lo cual parece ser poco frecuente en el habla puertorriqueña (Medina-Rivera 1997).

No hay una manera de comparar estos resultados con otros estudios, pero los mismos sirven para reafirmar la idea de que el fenómeno de lateralización en el español de Puerto Rico va más allá de los registros orales. Claramente, el fenómeno no se da en todos los hablantes, y se presenta con diferente intensidad como en el caso de Pablo y Francisco. No se realizaron entrevistas orales con los participantes, pero hay que mencionar que se escogieron a estos participantes para este estudio porque precisamente ambos mostraban notablemente casos de lateralización en la expresión oral.

Obviamente los resultados de este estudio no se elevan a los niveles por encima del 30% que se examinaron en la sección anterior pero sirven para ver un continuo de lateralización en hablantes puertorriqueños. Hay que recordar que las formas no estándar se pueden observar en los registros escritos, pero no en la misma intensidad que en los registros orales. También habría que tomar en cuenta el propósito y la oportunidad para una lengua menos cuidadosa (para abundar sobre este tema, se puede consultar Finegan & Biber 1994). Medina-Rivera (1997), en su estudio sobre el español de Puerto Rico entre jóvenes adultos de la ciudad de Caguas, toma en consideración registros diferentes al analizar el proceso de lateralización, tal como se presenta en la tabla 3:

Entrevista sociolingüística	52.9%
Conversación en grupo	62.8%
Presentación oral	6.6%

**Tabla 3.** Frecuencias de lateralización o variantes no estándar de /r/ en tres registros orales

Los datos de la tabla 3, tomados de Medina-Rivera (1997), agrupan los resultados no solo de la variante lateral sino que también los de otras dos variantes, elisión y aspiración de la vibrante simple, aunque estos no representan más del 5% de los casos de variantes no estándar. También hay que recordar que las tres frecuencias que aparecen en la Tabla 3 incluyen solamente los resultados de cuatro de los hablantes del estudio completo, los cuales se sometieron a los tres tipos de situaciones que se presentan en la Tabla 3. Eso explica la discrepancia entre la frecuencia de la entrevista en esta tabla y el valor de 30.8% que se ofreció en la Tabla 1. Seguramente los hablantes del estudio de 1997 no reflejarían casos de cambio de ‘r’ por ‘l’ en la escritura, ya que cuentan con un nivel educativo más alto que el de Pablo y Francisco, sin embargo ambos estudios nos permiten apreciar el fenómeno en hablantes con distintas características educativas, y la manera en la que un mismo fenómeno lingüístico se puede apreciar en registros orales y escritos.

#### 3.4.2. Manifestaciones en *Facebook*

Los medios cibernéticos con su cantidad de redes sociales le proveen al sociolingüista nuevas posibilidades para el análisis de la lengua y sus diferentes variaciones. El segundo componente del presente estudio es una muestra de ejemplos tomados de Facebook entre los años 2010-2013. A diferencia de la lengua hablada, en los registros escritos necesitamos determinar por el contexto, el tono o el nivel de seriedad de lo que se escribe. Es de conocimiento general las reacciones que pueden crear ciertos correos electrónicos o hasta cartas circulares de la administración, dependiendo del tono o actitud con las que se leen. En el habla oral podemos autocorregirnos, atenuar el tono, o hasta dar explicaciones inmediatas en caso de que haya confusión, en un registro escrito es mucho más complicado. Las caritas o emoticones son de gran utilidad a la hora de establecer dicho tono, ya que es muy fácil leer un mensaje con un tono o intención diferente del que originalmente supuso el que escribió el mensaje.

Cuando se escribe un mensaje en Facebook o en otro medio cibernético existe la posibilidad de cometer errores ortográficos, pero también existe la posibilidad de publicar un mensaje con un error claramente intencionado. La intencionalidad de un error en un mensaje puede tener una motivación lúdica, ya sea para estereotipar o burlarse de la manera de hablar de una persona o grupo dialectal, o ya sea para representar la manera “no estándar” de hablar de un grupo o individuo. David Crystal (2001: 38), en su libro sobre

la lengua y el internet, examina la importancia de las caritas o emoticones para establecer el nivel emotivo del mensaje, pero también señala que los “smileys have other roles than disambiguation. Sometimes they seem to be doing little more than expressing rapport. Often, their presence seems to have purely pragmatic force – acting as a warning to the recipient(s) that the sender is worried about the effect a sentence might have”. En este sentido el lenguaje del internet tiene elementos en común tanto con el lenguaje oral como el escrito, pues aparte de mostrar un mensaje escrito, muchas veces tiene la simplificación y el nivel emotivo del diálogo.

En inglés se pueden ver ejemplos como ‘dunno’, ‘wanna’, ‘ain’t’ los cuales ejemplifican formas del vernáculo que son de uso común y que se caracterizan por la simplificación de formas gramaticales. De manera similar en español es común ver formas como ‘pos’ por ‘pues’, ‘ta’ por ‘estar’ en las redes sociales, la cuales también sirven para representar la manera de hablar en un contexto no formal. Este tipo de ejemplos muestran el alto nivel de conciencia lingüística de los hablantes y su conocimiento sobre las formas dialectales que los caracterizan.

En el español de Puerto Rico se juega, hasta cierto punto, con las distinciones fonológicas que lo identifican. Hay gente que utiliza la variante velar de la vibrante múltiple para poner énfasis:

- (3) a. rico → [ríko]  
 b. arroz → [aróh]

Desde la década del 1980, por influencia de la televisión local, se puso de moda el piropo:

- (4) ¡Arroz, que carne hay! [aróh ke kálne ay]

Y se escuchaban las variantes [kálne], [káhne] o hasta una variante glotal [káʔne] en lugar de la vibrante simple o una variante aproximada. Los ejemplos 3 y 4 van más allá de representar una variante fonológica de uso en Puerto Rico y se convierten hasta cierta manera en rasgos fonológicos que caracterizan la identidad, que sirven como marcadores de un registro coloquial o que simplemente tienen una función lúdica. Valentín-Márquez (2007: 295) señala en su análisis de actitudes lingüísticas que “The participants on the Island, on the other hand, referred to lateralization of (r) the most to describe core Puerto Rican speech”. Los participantes de su estudio tienen conciencia del cambio de ‘r’ por ‘l’ y hasta lo ven como un elemento distintivo del habla puertorriqueña.

El aspecto lúdico de la lengua y la conexión nacional pueden relacionarse en el sentido de que dentro de un grupo étnico o cultural se crea humor con ciertos elementos lingüísticos. Pensemos solamente en el uso de la palabra *nigger* o *nigga* entre los afroamericanos de los Estados Unidos. La palabra tiene un estigma fuera del grupo de los afroamericanos, pero entre los afroamericanos utilizan la expresión como elemento de solidaridad y de humor (Spears 1998).

Dentro de la comunidad puertorriqueña, sobre todo la que vive en la Isla ya que tiene acceso a un mayor número de registros, se maneja el aspecto lúdico a nivel interno en el caso de la lateralización (o sus variantes: aspirada y glotal). Este juego que se crea con la lateralización se permite dentro del grupo, pero crea molestias o hasta discusiones cuando el juego viene de un hispanohablante que no es puertorriqueño. Uno de los programas radiales más populares de la Isla se llama ‘El Goldo y la Pelúa’. El mismo aparece como una réplica o parodia del famoso programa ‘El Gordo y la Flaca’ de la cadena televisiva Univisión. El programa de Univisión tiene como presentadores a dos cubanos, mientras que ‘El Goldo y la Pelúa’ hace referencia a un programa puertorriqueño, y la palabra ‘Goldo’ sirve para marcar esa identidad nacional. La palabra ‘Goldo’ marca además el tono del programa y del estilo de habla que manejan los locutores del mismo y muestra el nivel de conciencia lingüística que un hablante tiene de su propio dialecto (Mase 1999, Malaver 2002).

Los ejemplos que se presentarán a continuación están tomados de Facebook entre los años 2010-2013. En todos los casos tuve que monitorear los mensajes de un mismo participante para asegurarme de que el cambio de ‘r’ por ‘l’ en la escritura se debiera a:

1. una extensión del proceso de lateralización que se refleja en la escritura
2. una manera de jugar con el lenguaje para expresar énfasis y humor

Veamos algunos ejemplos:

- (5) En alanta esperando por mi buelo que salga para cleveland loco por estal con mi negrita y los nene los eztrano mucho depues de dos semana (Miguel, 4 junio 2013)
- (6) 100% orgulloso d ser boricua y loco pol volver..... aunque naciera en la Luna (Radamés 11 de septiembre 2010)

- (7) Bobo Tqm . Tas Pelddio , Yaa Nii Mee Ablas ! Tee toii Velanddo ! (Yessenia, 5 de noviembre 2011)
- (8) Tengo una listita a los cuales eso les debió haber pasado! Que fuellllte (Diana, 10 de agosto de 2013)
- (9) Lol, miraaaa que fuellllte!?!.. jajajaja (Diana, 20 de mayo 2013)
- (10) Ea visne! Pero si ya el carajito le gusta 'errrr plátano' lol (Diana, 23 mayo 2013)
- (11) Ufff, voy a ver si comienzo un fuego pa'que venga este y me apague la llamamcon la manguera! Jajajajaja, que fuelllllte, fuk it! (Diana, 18 junio 2013)
- (12) Primer dia del "kindel". Vieron mi bultito de Pacheco? (Cynthia, 19 mayo 2013)

Los ejemplos (5-7) representan casos en donde el fenómeno fonológico se transmite del nivel oral al nivel escrito, mientras que los ejemplos 6-12 son claramente lúdicos en su contexto.

El ejemplo (5) muestra varios errores ortográficos y representan la manera típica de esta persona escribir en Facebook. En este caso el ejemplo que es de interés para este estudio es el uso de 'estal' que concuerda con los resultados del estudio anterior en el cual se menciona que es precisamente en los verbos infinitivos en donde se muestra el cambio de 'r' por 'l' con mayor frecuencia. Cabe mencionar que Miguel es hermano de los sujetos Pablo y Francisco del estudio anterior, vive en Cleveland y se desenvuelve generalmente en español en su expresión oral y solamente en español en Facebook. Los ejemplos (6) y (7) son similares: en uno tenemos la palabra pol y en el otro pelddio. En el caso de pol parece representar la manera de hablar del participante y su manera de escribir, al igual que pelddio. La duplicación dd en peldio parece añadir énfasis a la expresión y no el cambio de 'r' por 'l'. Ambos participantes muestran también varios y constantes errores en su escritura. Son los más jóvenes de la muestra y su manera de escribir refleja lo que se mencionó anteriormente en el informe de la oficina del College Board y el Departamento de Educación. Al consultar este tipo de problemas ortográficos con maestras de español en la Isla, sugieren que se debe al cambio en el enfoque de la enseñanza del español, ya que antes el enfoque era mucho mayor en la gramática y ahora en la parte comunicativa, siguiendo el modelo de 'language arts' que se sigue en los Estados Unidos. Antes no se veía que a nivel de escuela secundaria se reflejaran las variantes fonológicas en la escritura, mientras que ahora parece ser común y frecuente.



Este es un tema de estudio que valdría la pena examinar entre estudiantes puertorriqueños.

Todos los ejemplos de Diana son claramente lúdicos: *fuelllte* en tres de los ejemplos y *errrrr*. Vale la pena volver a hacer referencia a la repetición del sonido que se quiere enfatizar, tal como se vio en el ejemplo (7). El ejemplo (10) muestra un caso de cambio de 'l' por 'r.' En términos generales la participante Diana escribe con claridad y con pocos errores ortográficos. Ella completó su educación universitaria en Puerto Rico en Lenguas Extranjeras, y puede comunicarse sin problemas tanto en español como en inglés. El contexto de cada uno de los mensajes es claramente jocoso, y de hecho una gran parte de sus mensajes en *Facebook* son de tipo humorístico. La palabra *fuelte* adquiere un significado equivalente a '¡Qué barbaridad!' o a algo que se sale de lo común y corriente. El mensaje está dirigido a los contactos puertorriqueños que tiene en *Facebook* y ellos reaccionan con risa no solamente al contenido, sino al uso de la palabra *fuelte*. Al observar otros mensajes de Diana en *Facebook*, no se encuentra un patrón de cambio de 'r' por 'l' en su escritura, pues en otros mensajes no se percibe de esa manera. Asimismo es interesante que sea precisamente la expresión *fuelte* la que se repita en tres de los mensajes, como si la palabra ya se hubiera gramaticalizado y adquirido un valor semántico diferente a la palabra *fuerte*.

El ejemplo de Cynthia en el (12) también es claramente lúdico, y en este caso pone la palabra entre comillas para poner el énfasis que quiere. Cynthia es maestra de inglés como segunda lengua en los Estados Unidos, aunque completó parte de sus estudios universitarios en Puerto Rico. Cynthia escribe en *Facebook* tanto en inglés como en español, y se mueve de lo formal a lo jocoso en sus temas. En este caso hace referencia a un retrato de sí misma cuando estaba en *kindergarten*. La representación de *kindel* enfatiza intencionalmente la manera de hablar de muchos puertorriqueños y su mensaje va dirigido a sus contactos puertorriqueños.

#### 4. CONCLUSIÓN

El fenómeno de lateralización en Puerto Rico va más allá de lo que los estudios muestran, ya que solo toman en consideración los registros orales. Las nuevas redes sociales de comunicación le brindan al sociolingüista otras fuentes para examinar el lenguaje y un mayor número de registros. Hasta cierto punto, los hablantes puertorriqueños utilizan formas no estándar como

el cambio de 'r' por 'l' como una manera de poner en evidencia un fenómeno que va adquiriendo un sentido de identidad nacional. El jugar con las propias distinciones fonológicas es una muestra de aceptar una característica que ha sufrido estigma en y fuera de la Isla.

Por otro lado, no podemos ignorar las observaciones del College Board presentadas en *El Nuevo Día*. Hay una preocupación válida de parte de las autoridades escolares de que los estudiantes puedan distinguir entre los registros orales y los escritos. De acuerdo al artículo del periódico (2006:7), la Facultad de Educación de la Universidad de Puerto Rico está tomando cartas en el asunto ya que de acuerdo a la Oficina del College Board el 40% de los nuevos estudiantes que ingresan en la universidad no cuenta con las destrezas necesarias y adecuadas para escribir español a nivel universitario.

Finalmente, los niveles de conciencia lingüística que se recogen en los ejemplos en los que se usa el lenguaje de una manera lúdica reflejan claramente la percepción que un hablante tiene de su dialecto (Presto 1999, Mase 1999, Malaver 2002). Con la cantidad de nuevas formas de obtener datos para un estudio lingüístico, es importante que los sociolingüistas vayan más allá de los métodos tradicionales para obtener sus muestras, y así contar con una idea más completa de la manera en la que los hablantes manipulan su dialecto y las maneras en las que reflejan las actitudes y percepciones sobre el mismo.



## REFERENCES

- Azevedo, Milton 2008 (1992): *Introducción a la lingüística española*, 3ra edición. Upper Saddle River, New Jersey: Prentice Hall.
- Crystal, David 2004 (2001): *Language and the Internet*, second edition. Cambridge: Cambridge University Press.
- El Nuevo Día* (2006): “Maestros con F en español en Puerto Rico”. En línea en: <http://www.puertorico.com/forums/history/21602-maestros-con-f-en-esoanol-en-page-ranking.html>.
- Emmanueli, Mirna (1986): *Actitudes lingüísticas hacia cuatro fenómenos Fonológicos*. Universidad de Puerto Rico, tesis de máster.
- Finegan, Edward y Douglas Biber (1994): “Register and social dialect variation: an integrated approach”, en Douglas Biber y Edward Finegan (eds.): *Sociolinguistic perspectives on register*. New York: Oxford University Press, pp. 325-347.
- Lambo, Edwin (2004): *Caribbean Spanish in the metropolis: Spanish language among Cubans, Dominicans, and Puerto Ricans in the New York City area*. New York: Routledge.
- López Morales, Humberto (1983): *Estratificación social del español de San Juan de Puerto Rico*. México, D.F.: Universidad Nacional Autónoma de México.
- Malaver, Irania (2002): “Dime cómo crees que hablas y te diré quién eres. Actitudes lingüísticas en la comunidad de habla caraqueña”, *Oralia* 5, pp. 181-201.
- Mase, Yoshio (1999): “Dialect consciousness and dialect divisions: Examples in the Nagano-Gifu boundary region”, en Dennis Preston (ed.): *Handbook of perceptual dialectology*. Amsterdam: John Benjamins, pp. 71-99.

- Medina-Rivera, Antonio (1997): *Phonological and Stylistic Variables in Puerto Rican Spanish*, University of Southern California, tesis doctoral.
- Niño Murcia, Mercedes (2011): “Variation and identity in the Americas”, en Manuel Díaz-Campos (ed.): *The handbook of Hispanic sociolinguistics*. West Sussex: Wiley-Blackwell, pp. 728-746.
- Preston, Dennis R. (1999): *Handbook of perceptual dialectology*. Amsterdam: John Benjamins.  
<http://dx.doi.org/10.1075/z.hpd1>
- Ramos-Pellicia, Michelle F. (2004): *Language contact and dialect contact: Cross generational phonological changes in a Puerto Rican community in the Midwest of the United States*, The Ohio State University, tesis doctoral.
- Simonet, Miquel, M. Rohena-Madrado y M. Paz (2008): “Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish”, en Laura Colantoni y Jeffrey Steele (eds.): *Selected Proceedings of the 3<sup>rd</sup> Conference on Laboratory approaches to Spanish phonology*. Somerville, MA: Cascadilla, pp. 72-86.
- Spears, Arthur K (1998): “African American language use: Ideology and so-called obscenity”, en Salikoko S. Mufwene, Guy Bailey, John Baugh y John R. Rickford (eds.): *African-American English*. New York: Routledge, pp. 226-250
- Valentín-Márquez, Wilfredo, 2007: *Doing being Boricua: Perception of national identity and the sociolinguistic distribution of liquid variables in Puerto Rican Spanish*, The University of Michigan, tesis doctoral.

## **El español de tierras altas y tierras bajas: sus reflejos en el español de Los Ángeles**

Claudia Parodi  
*Centro de Estudios del Español  
de los Estados Unidos (CEEUS)  
University of California, Los Angeles*

**Abstract:** Starting in the 1980s, Salvadorian and Guatemalan workers escaping political and economic instability in their respective countries have migrated in big waves to Los Angeles, California. This fact has originated a situation of dialects of Spanish in contact among the working class population, given that individuals coming from different geographical zones co-exist in the same area and share the same socioeconomic level. This article proposes to adopt the classification “Highland Spanish” and “Lowland Spanish” as a typology in order to analyze the changes that occur in the dialects of Spanish in contact. Speech samples were used from 120 first –and second– generation informants originally from high and low lands, and who participated in sociolinguistic interviews. These samples indicated that Mexican-based Highland Spanish is dominant in Los Angeles.

**Keywords:** Highland Spanish, Lowland Spanish, Dialects in contact, Spanish in the United States, Vernacular Spanish of Los Angeles (LAVS).



## 1. INTRODUCCIÓN

El estudio del español en contacto con el inglés en Los Ángeles ha avanzado notablemente gracias a los trabajos de Carmen Silva-Corvalán, quien ha analizado la permeabilidad de ambas lenguas y la correlación de factores lingüísticos y sociales que han motivado la variación y el cambio en contextos bilingües en trabajos tales como, entre otros, Silva-Corvalán (1986), (1994) y Silva-Corvalán y Gutiérrez (1995). Carmen Silva-Corvalán (1994) nos ha mostrado que, desde una perspectiva de la familia, para la tercera generación la mayor parte de los hispanos suelen perder el español o, a lo más, suelen tener una capacidad comunicativa muy limitada en esta lengua (Silva-Corvalán 1994). Lo mismo han encontrado Zentella (1998) en Nueva York y Veltman (1983, 1990) en la mayor parte de los hispanos de los Estados Unidos.

Históricamente, el español se ha hablado desde la época más temprana de la colonización española en el territorio norte de América, en lo que ahora es el suroeste de los Estados Unidos (Nuevo México, California, Arizona, Nevada, Utah, partes de Colorado, Wyoming, Oklahoma, Kansas y Texas) y en una porción de la Florida (San Agustín). Pero desde el tratado de Guadalupe Hidalgo de 1848, el suroeste de los Estados Unidos pasó a formar parte de la nación norteamericana. Con ello, surgieron varios cambios políticos y lingüísticos en esta zona. El inglés predominó sobre el español, aunque este último logró sobrevivir de manera subordinada al primero, en lo que Lope Blanch (2004: 154), ha llamado “español tradicional”, en poblaciones aisladas de Nuevo México y Colorado y que han estudiado desde Aurelio Macedonio Espinosa (1911) hasta Daniel Villa (2012) pasando por Garland Bills y Neddy Vigil (2008).

En el resto del suroeste, incluida California, señala Lope Blanch (2004: 154) que: “Las sucesivas oleadas de inmigrantes mexicanos a lo largo de nuestro siglo han ido llevando al suroeste estadounidense las modalidades lingüísticas propias del español de México [sobre todo de clase trabajadora<sup>1</sup>], *remexicanizando* en buena medida las hablas tradicionales de aquella amplia región”. Es decir que el español que encuentran los inmigrantes que llegan a Los Ángeles y que adquieren sus hijos, hablantes de español de segunda generación, es una variante del español mexicano, cuya “re-mexicanización”

---

<sup>1</sup> Utilizo el término clase trabajadora para referirme a personas de escasos recursos que proceden de zonas rurales de Latinoamérica o de suburbios urbanos donde estas se asientan, como ciudad Netzahualcóyotl en México.



resultó en una *koiné* de dialectos de México, la mayoría de clase trabajadora, procedentes en gran medida del Norte de México y del Bajío (Parodi 2011). Dicha *koiné* se formó sobre todo con el arribo de grandes oleadas de mexicanos a partir de la revolución de 1910 y quedó plenamente establecida para los años ochenta, en que llegaron subsiguientes inmigraciones masivas de trabajadores de habla española de otro origen distinto del mexicano, sobre todo centroamericanos, o mexicanos del sur de la república, como de Oaxaca, varios de ellos de origen indígena (Parodi 2013). Dado que me he referido a los rasgos más característicos de dicha *koiné* en otros trabajos (Parodi 2004, 2006, 2011), aquí solo menciono las peculiaridades de ciertos dialectos del español mexicano que se pierden en la *koiné* angelina.

Cabe añadir que tal *koiné* de dialectos del español mexicano popular hablada en su mayoría por personas de segunda generación de inmigrantes y algunos de generaciones posteriores, está formada por rasgos nivelados de dialectos del español mexicano. He denominado LAVS (por sus siglas en inglés<sup>2</sup>) a esta variante no estandarizada y oral del español vernáculo de Los Ángeles, la cual se usa tan solo en los barrios latinos de clase trabajadora y en zonas de contacto del español de esta misma clase socioeconómica. Además de los rasgos nivelados de base mexicana, LAVS se caracteriza por incorporar un gran número de préstamos del inglés y por el empleo frecuente del cambio de códigos (code-switching). Como se verá a lo largo de este trabajo, adquieren esta variante los hablantes de español que migran a Los Ángeles, incluidos aquellos que llegan de tierras bajas (sobre todo centroamericanos<sup>3</sup>). LAVS, como *koiné* que es, excluye los rasgos más marcados de los dialectos del español mexicano. En el español vernáculo de Los Ángeles no se aspira la /s/, como sucede en algunas zonas costeras de México (Moreno 1994, p. 78 y ss.), no suelen cerrarse las vocales medias en posición final de palabra, lo cual ocurre en ciertas zonas del Bajío (Moreno, 1994 p. 42 y ss), ni se fricativiza la /ç/, como sucede en el norte de México, incluida Baja California, y las costas de Jalisco, Colima y Michoacán (Moreno 1994, p.122). El vernáculo de Los Ángeles suele tener, además, un patrón de entonación peculiar (Andrade 2012). El uso del español en Los Ángeles no se limita al hogar, sino que también se emplea en contextos informales en la escuela, en la iglesia, en el trabajo, en los centros deportivos, en los mercados, en los comercios, en los

<sup>2</sup> Las siglas corresponden al inglés "Los Angeles Vernacular Spanish".

<sup>3</sup> En cuanto a los hablantes de tercera generación, aunque hablan poco el español, cuando lo hablan, también utilizan esta variante. Por lo regular, muestran una influencia del inglés mucho más fuerte que la segunda generación de inmigrantes.

restaurantes, etc. muchas veces junto con el inglés en los distintos barrios hispanos como, entre otros, “East Los Ángeles”, “Pico Rivera” y “Huntington Park” (Allen y Turner 2002). Por ejemplo, son muy frecuentes los anuncios comerciales en español vernáculo en distintos barrios de Los Ángeles (véase Carr 2013). Incluso es posible vivir en Los Ángeles sin saber el inglés.

A esta circunstancia cabe añadir, como ya mencionamos, que a partir de los años ochenta se han trasladado a Los Ángeles grandes oleadas de trabajadores centroamericanos, sobre todo salvadoreños y guatemaltecos a raíz de la situación política y económica de sus países, al igual que suramericanos y caribeños, aunque en menor medida (véase la tabla 3 abajo). Ello ha producido una situación de dialectos del español en contacto entre la población de clase trabajadora en Los Ángeles, pues se encuentran coexistiendo en una misma área geográfica —en la casa, en la escuela, en la iglesia, en el trabajo, etc.— y en un mismo nivel socioeconómico personas procedentes de distintas zonas geográficas. Los nuevos inmigrantes coexisten con personas que han conformado una comunidad de habla que emplea el “español vernáculo de Los Ángeles” o “español chicano de Los Ángeles” o LAVS.

La mayor parte de los trabajadores centroamericanos utilizan un español de *tierras bajas*<sup>4</sup>, en tanto que casi todos los trabajadores de base mexicana hablan LAVS, que es un español de *tierras altas*. Por esta razón, propongo que se estudie el contacto dialectal de las distintas variantes del español trasladadas a Los Ángeles usando el modelo de *tierras altas* y *tierras bajas*, mismo que permite analizar qué situaciones de contacto dialectal suelen darse entre los latinos<sup>5</sup> trasladados a Los Ángeles. Tras presentar tal modelo, ejemplifico los cambios que han surgido entre los hablantes de español salvadoreño (hablantes de español de tierras bajas) al entrar en contacto con los hablantes de español chicano o español vernáculo de Los

<sup>4</sup> No obstante que Guatemala sea zona de transición entre tierras altas y bajas, las zonas costeras de este país y de México (Tabasco y Chiapas), se acercan más a las tierras bajas de las áreas centroamericanas debido a que allí se aspira la /s/ en posición de coda, además de velarizarse la /n/ en posición final de palabra.

<sup>5</sup> En Los Ángeles este grupo étnico prefiere auto-identificarse como “latino”, “chicano”, “méxico-americano” o “mexicano”. El término “hispano”, utilizado por el censo, suele rechazarse ya que se considera “a white person’s word” (Fought, 2003, p.17). Nótese que el término “latino”, en esta zona ha cambiado de significado, pues se usa para referirse “to people of any Spanish-speaking background” (Fought 2003, p.17) o incluso para aludir solo a personas de origen mexicano. En este trabajo utilizo indistintamente español vernáculo de Los Ángeles (LAVS) y español chicano de Los Ángeles para distinguirlo del español hablado en otras zonas de los Estados Unidos. Para el uso de español chicano, véase Sánchez (1994) y para Chicano English, véase Fought (2003).

Ángeles (hablantes de español de tierras altas). Esta variante del español mexicano predomina en dicha área por las razones históricas y migratorias ya apuntadas. Goza de mayor prestigio que las otras variantes del español en la clase trabajadora y además sus hablantes tienden a imponer su forma de hablar a quienes emplean algún dialecto distinto del suyo<sup>6</sup>.

A la pregunta: “¿qué grupo de latinos te cae bien?”, una informante salvadoreña que llegó después de los trece años dijo preferir a los mexicanos y contestó: “quien me cae mejor o lo que sea, pues tengo más amigos mexicanos que cualquier otro tipo de latino” (MG/100B). Algunos hablantes salvadoreños abandonan rasgos como el voseo a raíz del contacto dialectal con los mexicanos —que son tuteantes— y porque tiene estigma en el propio El Salvador (Quesada 2000). Ambos factores motivan que desarrollen actitudes desfavorables hacia su propia habla original, especialmente frente al voseo: “no me gusta cómo hablamos nosotros los salvadoreños porque algunas personas se oyen bien maleducados...a mí no me gusta decir ‘¿Cómo estás vos?’ o ‘pasate vos’... digo ‘pasas adelante’ o ‘¿cómo estás tú?’”(AA /100B). Otros se acomodan a la variante mexicana rápidamente: “cuando llegué, lo tenía el acento, obviamente lo tenía bien marcado el acento salvadoreño y poco por poco me imagino porque mis amigos, la mayoría, desde que llegué, han sido mexicanos porque esa es la mayoría que hay aquí en Los Ángeles” (MG/100B).

En las últimas décadas se ha investigado el contacto dialectal de los hispanos en Estados Unidos desde varias perspectivas. Por ejemplo, Ana Celia Zentella, (1990), tras señalar la necesidad del estudio de los distintos dialectos del español trasladado a los Estados Unidos a raíz de la reciente diversidad lingüística causada por la migración, analiza el vocabulario de puertorriqueños, cubanos, dominicanos y colombianos en Nueva York. Nota que cuando hay variación léxica en los dialectos estudiados, los hablantes recurren con frecuencia al vocabulario inglés como una forma de “lexical leveling”. Pero en lo que atañe a la pronunciación, encuentra que el español de tierras altas o “conservative Spanish” se opone al español de tierras bajas o “radical Spanish”, que caracteriza al español caribeño. Anota que este último tiene estigma cuando se trata de puertorriqueños y dominicanos de clase trabajadora, pero no con los cubanos o colombianos costeros de clase media, a pesar de las semejanzas lingüísticas que tienen con los primeros.

<sup>6</sup> La forma de imponer el vernáculo de Los Ángeles frecuentemente es a través de la burla, sobre todo cuando los hablantes vosean. Lo mismo encuentran en Houston Aaron y Hernández (2007) por parte de los mexicanos cuando escuchan la aspiración de /s/ de los salvadoreños.

Esteban Hernández (2002), Hernández (2009) y Hernández y Maldonado (2012) reportan que los salvadoreños de primera generación se acomodan a sus interlocutores mexicanos en Texas (Houston y Brownsville), pues reducen tanto el voseo como la /s/ en posición de coda y la /n/ en posición final de palabra cuando dialogan con los mexicanos de la zona, que no vosean, no aspiran, ni velarizan la /n/. De igual modo, Parodi (2004) y Raymond (2012, 2012a) hallan que mientras las primeras generaciones de salvadoreños de clase trabajadora trasladados a Los Ángeles mantienen algunos de los rasgos de su dialecto original, sus hijos (o segunda generación) adquieren el español vernáculo de Los Ángeles y no el español salvadoreño. Potowski (2008, 2011) investiga el contacto dialectal en familias mixtas de mexicanos y puertorriqueños en Chicago. Muestra que en dicha ciudad, aunque predomina la variante materna, también se dan resultados mixtos. Frente al caso de Chicago, como ya indicamos, en Los Ángeles existe una variante del español de tierras altas, resultado de una koiné del español mexicano. Esta variante es un habla coloquial empleada sobre todo por la segunda generación de inmigrantes de clase trabajadora ha sido estable históricamente. En efecto, a partir de 1910 se han trasladado año con año a Los Ángeles —como puede verse en los censos— grandes contingentes de trabajadores mexicanos<sup>7</sup>. Por ello contamos con personas de segunda generación de clase trabajadora de todas las edades que han interactuado entre sí por más de un siglo. En Los Ángeles estos hablantes, además del entorno familiar, se comunican en español en contextos coloquiales entre amigos y dentro de sus comunidades en las tiendas, iglesias, clubes sociales, etc. empleando LAVS, la cual se ha difundido a lo largo Los Ángeles en la clase trabajadora. La inmigración constante facilita la manutención de esta comunidad de habla, lo cual a su vez hace innecesario contar con una tercera generación en ninguna familia para que se mantenga una variante dialectal. Como veremos más adelante, es suficiente que haya continuidad de segundas generaciones a lo largo del tiempo para que se forme una koiné<sup>8</sup>. En Nuevo México, en cambio, Bills y Vigil (2008) registran dos variantes. La primera es un español tradicional, hablado por los norteños de la zona, seguidores de la tradición colonial en

<sup>7</sup> Obviamente el español mexicano, como el español de cualquier otra región, ofrece un buen número de variantes dialectales, las cuales se abandonan a favor de LAVS en Los Ángeles. Para las variantes del español mexicano, véase Moreno de Alba, (1994).

<sup>8</sup> No es necesario que haya terceras generaciones de inmigrantes para que se forme una koiné como ha sugerido Kerswill, (2000: 670), quien admite que puede darse una koiné en una generación.

el norte de Nuevo México y el sur de Colorado. La segunda variante es un español más moderno, hablado por los sureños de Nuevo México, similar al español de los pueblos del norte de México.

En el sur de la Florida, sobre todo en Miami, prevalece un español de base cubana de clase media, cuyos hablantes llegaron poco después del gobierno de Fidel Castro. Sin embargo, esta variante se niveló con el español de clase trabajadora que trajeron a los Estados Unidos los marielitos cubanos años después (Lipski 2008). Ricardo Otheguy y Ana Celia Zentella (2012) encuentran que en Nueva York hay cierta preferencia por el abandono de la -s aspirada en hablantes caribeños a favor de la -s no aspirada a raíz del contacto dialectal. Pero en lo que atañe al uso del pronombre sujeto hallan que el patrón caribeño (Puerto Rico, Cuba y Santo Domingo) predomina sobre el uso continental (Colombia, Ecuador y México) en la ciudad mencionada. Bayley y otros (2012) señalan que el uso de pronombres sujeto en San Antonio Texas entre puertorriqueños y mexicanos depende de las redes sociales y profesionales. Los hablantes que se relacionan con puertorriqueños emplean los pronombres sujeto de manera similar a San Juan o Nueva York, pero los hablantes que se relacionan sobre todo con personas de origen mexicano, utilizan el pronombre sujeto de manera semejante a los últimos. Otheguy, Zentella y Livert (2007) apuntan que en Nueva York el contacto del español con el inglés, junto con el contacto de dialectos del español, motivan que se forme una comunidad de habla del español neoyorquino entre la segunda generación de migrantes latinos.

Como puede observarse en los estudios de contacto de dialectos del español mencionados, los distintos grupos de latinos arraigados en los Estados Unidos presentan diferencias regionales que caracterizan el español de dicho país. En tanto que en Los Ángeles predomina LAVS, una variante del español de México, en ciertas áreas hay variación como en San Antonio, Texas, o en Chicago, pero en otros lugares del propio Texas, como Houston o Brownsville, también predomina el español mexicano. En cambio, en Nueva York se empieza a formar una comunidad de habla de español neoyorquino, resultado de la nivelación de dialectos caribeños y del español continental entre los hablantes de segunda generación. En Nuevo México, en cambio, se registran dos variantes, el español tradicional y el español moderno de clase trabajadora. En Miami prevalece un español de base cubana.

A pesar de que las terceras generaciones de inmigrantes tienden a perder su lengua étnica, el caso del español de Los Ángeles resulta particularmente interesante pues, como ya señalamos, dada la tradición

histórica y continuidad migratoria de hispanohablantes de clase trabajadora procedentes de México, han convivido hablantes de segunda generación de todas las edades desde hace más de cien años en esta zona. Ello ha permitido que se formara la koiné de dialectos de base mexicana que dio origen al vernáculo de Los Ángeles, gracias al contacto intra-grupal en la comunidad de habla hispanohablante de clase trabajadora. Dicho vernáculo, como ya vimos, es una variante del español de tierras altas de base mexicana que elimina rasgos dialectales como la nasal epentética al final de una palabra terminada en –s y el cierre de vocales medias en posición final, no aspira –s, incorpora préstamos del inglés, selecciona un vocabulario específico y tiene una entonación particular. Sus hablantes, además, utilizan con frecuencia el cambio de códigos.

## **2. EL ESPAÑOL DE TIERRAS ALTAS Y TIERRAS BAJAS: UN MODELO TIPOLÓGICO DEL ESPAÑOL LATINOAMERICANO**

No obstante que el español se habla en diecinueve países de América Latina y en gran parte de los Estados Unidos, si se dejan de lado las peculiaridades lingüísticas de cada lugar y si se atiende a ciertos rasgos generales de la pronunciación, resulta posible hacer una abstracción y dividir el español americano en dos grandes tipos lingüísticos: español de tierras altas y español de tierras bajas. El primero, aunque es seseante y yeísta, se aproxima al español septentrional (Castilla la Nueva, Castilla la Vieja) y se habla sobre todo en las tierras altas o zona central del continente. En cambio el segundo, también seseante y yeísta, se asemeja al español meridional de España (Extremadura, Andalucía y Canarias) y se habla en las costas de América, sobre todo las del Caribe y en Centroamérica, excluida Guatemala, la cual tiene rasgos de tierras altas y bajas, como puede observarse en los mapas fonéticos que proporciona Canfield (1981). El origen de esta división tipológica se debe a Pedro Henríquez Ureña (1921) y la retomaron Menéndez Pidal (1962) y Ángel Rosenblat (1967). El primero reinterpreta las tierras bajas como zonas frecuentadas por la flota española durante la colonia y las tierras altas como regiones del interior, en tanto que el segundo autor mantiene la terminología de Henríquez Ureña, misma que prevalece entre otros estudiosos del español americano. Menéndez Pidal (1962: 142) atribuye los inicios de esta división a razones históricas de la colonización, pues:

las costas frecuentadas por la navegación de las flotas [procedentes de Sevilla, Cádiz y Canarias] recibieron más directa, íntima y persistentemente las nuevas ondas lingüísticas del habla familiar metropolitana, en contraste con las regiones del interior.

Entre otros ejemplos, Menéndez Pidal cita la ciudad de México como zona del interior (o alta) frente a Veracruz, Tabasco y Acapulco como zonas de las flotas (o bajas). En Colombia, distingue Bogotá, como la primera, (tierras altas) y el interior de los puertos de Santa Marta y Cartagena, como la segunda (tierras bajas). En Ecuador, diferencia la zona interior o alta de Quito en oposición a las tierras de las flotas o bajas de Puerto Viejo y Guayaquil. Rosenblat (1967: 110) define “tierras bajas” como “las vastas regiones de las costas y los llanos que las prolongan”, y “tierras altas” como “la vasta porción de las mesetas que se extienden desde Méjico a través de la cordillera de América central y de los Andes de América del Sur hasta el norte argentino”. Los términos tierras altas y tierras bajas obedecen a una generalización que abarca un buen número de zonas geográficas de América Latina, pero tal división no debe tomarse ni de manera literal, ni en un sentido climatológico, sino como una tipología de gran utilidad para entender el contacto dialectal y las características más relevantes del español de América. Esta tipología del español americano podría reinterpretarse de varias maneras, por ejemplo, como español conservador, el de tierras altas, y español innovador, el de tierras bajas.

Las peculiaridades fonéticas que corresponden a una y otra variante las he utilizado como clasificación criterio tipológico para identificar los cambios que ocurren en el español de los Estados Unidos, especialmente en Los Ángeles, California, empleando los métodos de la sociolingüística urbana. Como puede observarse en el cuadro 3, en Los Ángeles emigran y se interrelacionan grandes contingentes de personas hispanohablantes que proceden de toda América y hablan distintos dialectos del español. Tales dialectos pueden analizarse a la luz de dicha tipología a fin de explicar los cambios fonéticos y fonológicos más significativos que surgen por el contacto dialectal. En efecto, a raíz del contacto de dialectos del español, sobre todo entre personas de clase trabajadora y sus descendientes —mismos que suelen interactuar estrechamente en varios contextos a los cuales me referiré en breve—, surgen cambios de gran interés para el conocimiento del contacto de dialectos en general y del español angelino en particular. Estos cambios



se originan de lo que he llamado “contacto intra-grupal”, pues afectan a los hablantes que conforman un mismo grupo sociolingüístico. El empleo de esta tipología permite distinguir, además, los cambios que se derivan del contacto de dialectos (variantes del español) de los cambios que surgen por el contacto de lenguas (en este caso el inglés y el español) entre los hablantes del español de los Estados Unidos. Los siguientes rasgos fonéticos caracterizan a uno y otro tipo lingüístico:

<i>Español de tierras bajas</i>	<i>Español de tierras altas</i>
1. Aspiración de /s/ s → h/o / __C, __# <b>[páh.ta], [ká.sah]</b>	no se aspira <b>[pás.ta] [ká.sas]</b> <pasta>, <casas>
2. Aspiración de /x/ x → h <b>[ká.ha]</b>	no se aspira <b>[ká.xa]</b> <caja>
3. Velarización de /n/ n → ŋ / __# <b>[páj]</b>	no se velariza <b>[pán]</b> <pan>
4. Pérdida de /d/ d → o / V__V <b>[kan.sá.a], [kan.sá]</b>	no hay pérdida <b>[kan.sá.áa]</b> <cansada>
5. Neutralización de /r/ y /l/ r → l / __C, # <b>[kál.ne], [ko.mél]</b>	no se neutraliza <b>[kár.ne], [ko.mér]</b> <comer>
6. Vocales átonas V-átonas mantenidas <b>[án.tes]</b>	V-átonas pueden debilitarse, o perderse <b>[ánt's]</b> <antes>

**Tabla 1.** Español de tierras altas y tierras bajas

Claro está que dichos rasgos no necesariamente han de encontrarse todos presentes en un dialecto del español para que se clasifique como perteneciente a una u otra tipología. Basta que en un dialecto estén presentes algunas de estas características de manera predominante. Las variables más representativas son la presencia o ausencia de aspiración y pérdida de /s/, aspiración de /x/ y la velarización o no velarización de /n/. Incluso, tomando



en cuenta sólo la aspiración, Raúl Ávila (2009) clasifica las variantes del español estándar en alfa, que no aspira: [pás.tas] y beta que sí aspira [páh.tah] frente a la peninsular septentrional o castellana, gama, que no aspira y distingue las sibilantes interdental /ə/ y la apicoalveolar /s/, como en *cielos* [əjé.loś] <cielos>. Cabe añadir a la tipología de tierras altas y bajas el *voseo*, otro rasgo de importancia para caracterizar el español de los Estados Unidos. El voseo no tiene relación con el español de tierras altas y bajas, pero en ciertos casos, como en el español de Los Ángeles, resulta de gran importancia para entender el contacto dialectal. El voseo es un rasgo característico y dominante en el español de América pero suele ignorarse dado que en muchos lugares no goza de prestigio. No se enseña en las escuelas y suele censurarse, incluso entre los propios hablantes que lo usan, como ya indicamos y volveremos a ello más adelante. Geográficamente, el voseo se encuentra extendido, pero no generalizado, en el centro y el sur América<sup>9</sup>. Está prácticamente extinto en el Norte de América y en Las Antillas. De hecho, en México no se utiliza, salvo en Tabasco y Chiapas, que en cierto momento formaron parte de Centroamérica<sup>10</sup>.

### 3. EL ESPAÑOL DE TIERRAS ALTAS Y TIERRAS BAJAS EN LOS ÁNGELES

El español vernáculo o español chicano de Los Ángeles es un dialecto de tierras altas, frente a otros dialectos trasladados a California como el salvadoreño, el hondureño o el cubano, que son de tierras bajas. El primero tiene características dialectales —léxicas, fonológicas y morfológicas— que los hablantes de otras variantes del español sobre todo los de clase trabajadora reconocen y adoptan en Los Ángeles por tratarse de la variante del español que predomina, tiene prestigio y tradición histórica en esta área, como pudo observarse en el análisis de 120 muestras de habla de entre 45 y 60 minutos, producidas en Los Ángeles, California, por informantes de clase trabajadora, distribuidos de según se muestra en la Tabla 2:

<sup>9</sup> El voseo se ha generalizado en Argentina, Uruguay y Paraguay pero mantiene el estigma en el resto de Suramérica. Ver, entre otros, Páez (1981), Moreno de Alba (2001), Lipski (1994), Quesada Pacheco (2000).

<sup>10</sup> Existe, además, el *ustedeo*, que es el uso del pronombre *usted* en alternancia con *tú* y a veces con *vos* como forma de confianza en ciertas regiones de Costa Rica, Panamá, Venezuela y Colombia. También se utiliza en los estratos sociales altos de Chile y Argentina. Hay poca información sobre el *ustedeo*. Algunos datos se encuentran para Centroamérica en Quesada Pacheco (2000) y para Centroamérica y Suramérica en Lipski (1994), donde se citan referencias.

edad de arribo a Los Ángeles	tierras altas, hombres	tierras altas, mujeres	tierras bajas, hombres	tierras bajas, mujeres
< 8 años o nacidos en LA	10	10	10	10
emigrados de 8 a 21 años	10	10	10	10
emigrados de 22 años en adelante	10	10	10	10
total	30	30	30	30

**Tabla 2.** Distribución de informantes

La mayoría de los informantes de tierras altas son de origen mexicano, sobre todo del Bajío, Zacatecas y el norte de México (58), y otros de Colombia (2). Los de tierras bajas son en su mayoría de El Salvador (46), Honduras (5) Panamá (4) Nicaragua (3) y Costa Rica (2)<sup>11</sup>. Esta selección de informantes se basa en los resultados de un estudio previo de la comunidad latina de Los Ángeles (Parodi 2004) en el cual pude observar que la edad juega el mismo papel en la adquisición de un dialecto que en una lengua. En efecto, en dicha investigación comprobé que la edad clave para adquirir un dialecto con la fluidez de un nativo es desde el nacimiento hasta los ocho años<sup>12</sup>. Después de esta edad, la adquisición es variable. Por ello, a fin de controlar la variación, en este trabajo dejé de lado a los informantes cuya edad fluctuara entre 4 y 13 años a fin de evitar cualquier tipo de ambigüedad que pudiera deberse a la edad, ya que quería probar si la división de dialectos de tierras altas y bajas era efectiva para caracterizar el español de Los Ángeles. Las

<sup>11</sup> Las muestras proceden de una base de datos del español de Los Ángeles que he venido reuniendo desde hace más de diez años. Hasta la fecha contamos con más de 750 audio grabaciones y video grabaciones. Se han transcrito diez páginas mecanografiadas de cada una de estas grabaciones.

<sup>12</sup> La correlación entre edad y adquisición de una lengua o un dialecto se ha propuesto desde varias perspectivas. Por un lado en la Gramática Generativa se habla de un "periodo crítico" (poco antes de la pubertad) y de innatismo para la adquisición de una lengua (Chomsky 1986). Por otro lado, estudios sociolingüísticos como, entre otros, Chambers (1992), Kerswill (1996), Parodi (2004) y Hernández (2009) corroboran la relevancia del factor edad para la adquisición de un dialecto. Chambers (1992) y Parodi (2004) encontraron que la edad para adquirir un dialecto de forma nativa es hasta el fin de los siete años. Parodi señala que quien aprende un dialecto, además de tener la edad apropiada, debe mostrar una actitud positiva hacia dicho dialecto y encontrarse en contacto continuo con compañeros (*peers*) de la misma edad que hablen la variedad en cuestión.

entrevistas que elegí proceden de la base de datos mencionada en la nota 7. La variación a que están sujetos los informantes puede verse en Parodi (2004)<sup>13</sup>. Las conversaciones se llevaron a cabo siguiendo un cuestionario sociolingüístico que contiene preguntas abiertas sobre la vida y opiniones de cada informante<sup>14</sup>. También utilicé diez videos de habla espontánea con dos o tres informantes dialogando libremente.

Analiqué las entrevistas tomando en cuenta los rasgos característicos de los dialectos del español americano de tierras altas y bajas a fin de determinar si los informantes de la primera generación hablaban un dialecto de El Salvador o de otro lugar de Centroamérica (de tierras bajas) o si hablaban español vernáculo de Los Ángeles o español chicano (de tierras altas). Para los hablantes de segunda generación, es decir personas nacidas en Los Ángeles, determiné si hablaban un dialecto de tierras bajas de Centroamérica (El Salvador, Honduras, Panamá, Nicaragua o Costa Rica), si hablaban español vernáculo de Los Ángeles, que es español de tierras altas o si eran bi-dialectales por utilizar ambas variantes en distintos contextos. Además pregunté directamente el significado de veinticinco regionalismos de El Salvador a fin de comprobar hasta qué punto los hablantes de otros dialectos conocían tales voces<sup>15</sup> y poder diagnosticar si hay permeabilidad de dialectos del español hablado en el sur de California. El foco central de esta investigación es mostrar las peculiaridades tipológicas principales del español vernáculo de Los Ángeles entre hablantes de primera generación de clase trabajadora y de segunda generación que se criaron en barrios de clase trabajadora, como por ejemplo el Este de Los Ángeles. Una de las principales preguntas era determinar si los hijos de los hablantes de español de tierras bajas de clase trabajadora adquieren el español de tierras bajas o el vernáculo de Los Ángeles (de tierras altas), a pesar de que sus padres hablen una variante de español (de tierras bajas), distinta de la vernácula de Los Ángeles.

---

<sup>13</sup> Este grupo de informantes resulta de mucho interés, pero dado su carácter mixto requiere de un estudio amplio de sus rasgos lingüísticos para poder hacer generalizaciones. Dejo esta tarea para investigaciones futuras.

<sup>14</sup> El cuestionario empleado puede verse en el apéndice 1 anexo a este trabajo.

<sup>15</sup> Determiné las palabras por investigar solicitando a varios hablantes salvadoreños listas de voces características de su país de origen. Después cotejé dichas palabras en Geoffroy Rivas (1978), Matías Romero (2005), y otros diccionarios como el DRAE. Aunque las palabras de la lista son conocidas en otras regiones, en el habla popular salvadoreña son de uso cotidiano. Además es importante señalar que no forman parte del vocabulario del español vernáculo de Los Ángeles. A cada informante se le preguntó: *¿Qué quiere decir X?* La lista de palabras puede verse en el apéndice 2 al final de este artículo.

Basta consultar los datos del censo del año 2000 para darse cuenta del predominio numérico de mexicanos en los condados de Los Ángeles, Riverside, Orange, San Bernardino y Ventura, los cuales forman la unidad estadística CMSA (*Consolidated Metropolitan Statistical Area*). En efecto, en el año 2000, de un total de 6.5 millones de hispanos o latinos, 4.5 millones eran de origen mexicano en dicha década. En total, los hispanos o latinos formaban el 40% de la población de esta área. Incluyo los datos del censo del año 2000, tomados de Allen y Turner (2002) relativos a la distribución de hispanos o latinos en dicha unidad estadística, en la Tabla 3.

Mexicanos	4,964,000	76.00%
Salvadoreños	213,000	3.30%
Guatemaltecos	118,000	1.80%
Suramericanos	109,000	1.70%
Otros centroamericanos	106,000	1.60%
Puertorriqueños	66,000	1.00%
Cubanos	54,000	0.83%
Otros hispanos	900,000	13.00%
<i>Total</i>	<i>6,565,000</i>	<i>99.20%</i>

**Tabla 3.** Población latina en Los Ángeles, año 2000

La hegemonía del español vernáculo de Los Ángeles en la comunidad trabajadora de este condado va ligada al predominio numérico de mexicanos que aumenta año con año debido a la inmigración de trabajadores de México, a la continuidad histórica de la migración de estos en el suroeste de los Estados Unidos desde principios del siglo xx y al hecho de que California fue parte del territorio que ahora es México durante la colonia (Sánchez 1994, Lamar Prieto 2012).

Cuando los adultos nacidos en distintos países latinoamericanos se encuentran en la zona de Los Ángeles, empiezan a agregar e incluso a sustituir elementos léxicos característicos de su habla regional por los términos más comunes del vernáculo californiano cuando son de clase trabajadora y viven en barrios en los que predominan personas de origen mexicano. Ello no sólo

se debe a un deseo de facilitar la comunicación, sino a evitar el estigma por parte de la población chicana y mexicana dominante que no suele tolerar con facilidad las diferencias dialectales. Los veinte salvadoreños nacidos en Centroamérica (100% de la 1ª generación de tierras bajas) hacen las siguientes sustituciones y ampliaciones a su vocabulario incluidas en la Tabla 3:

<i>salvadoreño</i>	<i>vernáculo de Los Ángeles</i>
cipote	muchacho, chamaco
pacha	botella, mamila, biberón
bayunco	bromista, simple
ayote	calabacita
chele	güero 'rubio'
chacalines	camarones

**Tabla 4.** Vocabulario salvadoreño y vocabulario chicano

Las sustituciones son momentáneas entre los hablantes emigrados a Los Ángeles desde los veintidós años en adelante, como resultado de la acomodación al interlocutor en el sentido de Giles y Coupeland (1991). Pero son permanentes en los veinte informantes de 2ª generación de tierras bajas (100% de los informantes del grupo que nacieron en LA o llegaron antes de los ocho años). Tales informantes utilizan el léxico del vernáculo angelino, pero tienen un conocimiento pasivo de las voces salvadoreñas, si son de origen salvadoreño (doce informantes 60%). Los otros ocho centroamericanos de segunda generación (40%) conocieron algunas de estas palabras (*cipote*, *pacha* y *chele*). Los informantes de origen mexicano y colombiano las desconocieron totalmente (sesenta informantes, 100%). Esta muestra indica que no hay expansión de las variantes salvadoreñas en la comunidad latina de clase trabajadora, a pesar de que las *pupusas* y el *curtido* son bien conocidos<sup>16</sup>.

En lo que atañe al contacto de sistemas fonéticos y fonológicos distintos, la sustitución de un sistema por otro, aunque sea parcial, no se da con tanta facilidad y rapidez como sucede con el léxico entre los adultos

<sup>16</sup> Las *pupusas* son una especie de crepas de maíz rellenas de queso, frijoles o chicharrón, similares a las "gorditas" mexicanas. El *curtido* acompaña a las pupusas. Se prepara con repollo o col, cebolla, vinagre, chile y otros condimentos.

hablantes de diversos dialectos regionales. Las entrevistas analizadas muestran que por lo regular los individuos de primera generación de tierras bajas que llegaron después de los veintidós años y que entraron en contacto con otro dialecto distinto del suyo después de la pubertad (español de tierras altas) no abandonan sus hábitos articulatorios originales, pues es difícil que adquieran cabalmente otro sistema fonético o fonológico (veinte informantes procedentes de tierras bajas, 100%). Pero los que llegaron entre catorce y veintidós años tienen rasgos de tierras altas y bajas (catorce de El Salvador 70%, dos de Honduras 10%, dos de Nicaragua 10%, uno de Panamá 5% y uno de Costa Rica 5%). Es decir, son bi-dialectales. Entre los mexicanos se dieron ciertas adaptaciones al vernáculo de Los Ángeles. Por ejemplo, los hablantes de Michoacán trasladados a la ciudad angelina (15 informantes de cuarenta informantes de primera generación, 37.5%) tienden a cerrar menos o no cerrar las vocales medias en posición final que los hablantes de Michoacán que permanecen en México, los cuales cierran las vocales con gran frecuencia (Moreno de Alba 1994, p. 45). Algunos ejemplos de cierre vocálico son: [lé.çi] por <leche> y [ka.βá.yu] por <caballo>. Asimismo, eliminan la nasal epentética después de /s/ final en ejemplos como [pwés'n] también característica de Michoacán y Jalisco (Moreno 1994, p. 125), que se realiza como [pwés] <pués> en el vernáculo de Los Ángeles.

Como puede verse, los inmigrantes que se trasladaron de adultos al sur de California mantienen sus sistemas fonológicos originales, aspiración de /s/, velarización de /n/, etc. En cambio los hijos de inmigrantes o segunda generación, que nacieron en Los Ángeles o los inmigrantes de primera generación que llegaron a este lugar antes de los ocho años, suelen adquirir el español chicano o vernáculo de los Ángeles, aunque sus padres hablen otro dialecto, siempre y cuando hayan vivido en áreas en que predominan los chicanos de clase trabajadora. Ello es común en Los Ángeles en el caso de los guatemaltecos y los salvadoreños de clase trabajadora en virtud de que suelen convivir estrechamente con los mexicanos, también de clase trabajadora (Parodi 2004). En cambio, los argentinos, los chilenos, algunos cubanos y varios puertorriqueños prefieren vivir en comunidades distintas de las chicanas de clase trabajadora, sobre todo cuando son de clase media. Las clases medias suelen mantener su identidad dialectal con mayor fuerza que la clase trabajadora. Buzatu (2012) reporta que los argentinos de clase media nacidos en Los Ángeles se distinguen de los hablantes de LAVS por aspirar la /s/ implosiva entre el 40% y el 70% de los casos.

En el caso de los salvadoreños de clase trabajadora, que son hablantes de español de tierras bajas, los rasgos fonológicos que éstos suelen cambiar al entrar en contacto con el español chicano (de tierras altas) son los siguientes incluidos en la Tabla 6:

/s/ no se aspira en /\_\_C, \_\_#  
**[páh.ta], [ká.sah] → [pást.a] [ká.sas]**

/x/ no se aspira  
**ká.ha] → [ká.xa]**

/n/ no se velariza en /\_\_C, \_\_#  
**[páj] → [pán]**

no hay epéntesis entre / í\_\_a  
**[dí.ya] → [dí.a]**

No se pierde /y/ entre / í,é\_\_V  
**[é.a] → [é.ja]**

**Tabla 6.** Rasgos fonéticos del español de tierras altas en hablantes de origen salvadoreño (de tierras bajas), 2ª generación.

Lo mismo encontró Raymond, 2012, en su estudio monográfico de una familia salvadoreña de cinco miembros trasladada a Los Ángeles: los padres emigraron después de los veintidós años y sus tres hijos nacieron en Los Ángeles. En tanto que los padres mantienen su diáspora salvadoreña, los hijos de veintidós, dieciocho y catorce años, utilizan el español vernáculo de Los Ángeles. La familia vive en Costa Mesa, donde el 75% de la población es de trabajadores de origen mexicano y un 8% es de salvadoreños, también de clase trabajadora. El porcentaje restante procede de otros grupos no especificados por el censo. Los cinco miembros de esta familia trabajan, estudian e interactúan con la comunidad mexicana a la cual se han integrado totalmente, a tal punto que muchas veces los han llamado “mexicanos”. Frente a las ciento veinte encuestas realizadas por múltiples encuestadores, que recogieron los materiales que usé para este estudio, Raymond grabó a la familia durante la cena por aproximadamente ciento veinte minutos en varias ocasiones por medio de audio grabaciones y video grabaciones.

Con respecto a la morfología, los tratamientos personales, como el tuteo o el voseo, que tienen implicaciones en el sistema verbal y pronominal, son

indicadores importantes del origen regional de los hablantes. Los hablantes de español vernáculo de Los Ángeles no son voseantes como tampoco lo son ni los caribeños ni los mexicanos (salvo los de Tabasco y Chiapas), a pesar de que la mayoría de los dialectos americanos son voseantes. Pero los hablantes que utilizan el *vos* en su dialecto original, como los salvadoreños, guatemaltecos y costarricenses, que se trasladaron después de la pubertad al sur de California, aprenden las equivalencias necesarias para manejar el *tuteo* junto con el *voseo*<sup>17</sup>, los cuales utilizan en contextos diferentes. En general vosean con sus familiares y sus compatriotas, pero tutean en los demás casos. Una informante salvadoreña que llegó a los catorce años a Los Ángeles indicó: “el *vos* trato de no usarlo porque yo sé que alguna gente ni sabe qué es el *vos* (MG/100B).

Los hijos de los salvadoreños y demás entrevistados centroamericanos nacidos en California no vosean, pues usan el español vernáculo de Los Ángeles. Pero en otras comunidades fuera de California las generaciones jóvenes mantienen el voseo (Woods y Rivera Mills 2012). A continuación, en la Tabla 7, presento algunos ejemplos que ilustran los tratamientos en el español de Los Ángeles tomados de entrevistas realizadas en Los Ángeles:

**Voseo en hablantes de 1ª generación (procedentes de El Salvador):**

- “*Tenés que patear el clutch*” (a un amigo salvadoreño)  
(ML/100B)
- “*Salite vos de mi carro*” (a un niño salvadoreño,  
en discurso reportado)  
(ML/100B)

**Tuteo en hablantes chicanos y en hablantes salvadoreños de 2ª generación:**

- “*No te pongas el saco*” (a un amigo chicano)  
(HP/100B)
- “*Tienes mucha suerte, Pepe*” (al hijo de un amigo chicano)  
(HP/100B)

<sup>17</sup> Como arriba indiqué, el voseo se ha definido como el uso sistemático del pronombre *vos* y sus correspondientes formas verbales en el tratamiento de la segunda persona del singular (Moreno de Alba 2001). En algunos países como Argentina está casi totalmente generalizado (no se vosea en la Patagonia ni en la Tierra de Fuego), pero en otros países alterna con el tuteo y suele estigmatizarse. Toda Centroamérica vosea, pero el tuteo tiene más prestigio (Quesada Pacheco 2000). El voseo presenta mucha variación geográfica y morfológica a lo largo del continente, como han mostrado Páez (1981) y Moreno de Alba (2001).



**Usted, generalizado en chicanos y salvadoreños, usado para marcar diferencias de rango social o cariño:**

“Cuántos años tenía <i>usted</i> ?”	(a una persona de respeto, en este caso la mamá)
	(AA/100B)
“Véngase, mi bichito, con su mamá”	(la mamá a su hijo)
	(AA/100B) <sup>18</sup>

**Tabla 7.** Tratamientos en Los Ángeles

Dado que la mayor parte de la población latina del sur de California es de origen rural, cuando hay contacto de dialectos, los hablantes suelen conservar los rasgos rurales comunes, los cuales son generales de toda el habla española de las aldeas y de las áreas populares de las ciudades en ambos lados del Atlántico. Por ello, el español vernáculo de Los Ángeles conserva una buena dosis de rasgos rurales, que cabe llamar rur-urbanos, los cuales suelen estigmatizarse entre los hablantes monolingües de cualquier variedad de español estándar y en los medios académicos de California, donde se enseña español estándar mexicano monolingüe (Parodi 2008). A pesar de ello, el español chicano cuenta con prestigio en la comunidad latina de clase trabajadora en Los Ángeles, como se indicó arriba. A continuación presento una lista de los rasgos de origen rur-urbano que se utilizan en el español de Los Ángeles (tomada de Parodi 2011):

**Algunas características de origen rur-urbano del español chicano**

1. Generalización del morfema -s de 2ª persona en las formas del pretérito simple:  
*comistes, llegastes, escribistes*
2. Terminación de la 2ª persona del imperfecto en -nos (verbos en -ar):  
*llegábanos, terminábanos, hablabábanos*
3. Contracción del artículo ante vocal<sup>19</sup>:  
*l'avena, l'alfalfa, l'espada*
4. Acentuación analógica en la segunda persona del presente de subjuntivo :  
*sépanos-sépanos, cómanos-cómanos, díganos-díganos*

<sup>18</sup> No obstante que encontré algunos casos esporádicos de *ustedeo* (uso de *usted* por *tú*) en alternancia con *tú* y *vos* entre colombianos y costarricenses que emigraron de adultos a Los Ángeles, estos fueron muy pocos para sacar generalizaciones sobre su uso, en especial porque llevaban menos de diez años en Los Ángeles.

<sup>19</sup> Las formas precedidas de asterisco también se usan en el habla estándar coloquial mexicana en habla rápida.

5. Pronunciación del fonema /f/ como /x/ ante el diptongo ue y en otros arcaísmos:  
     /*xwe/ por /fue/, /xwerte/ por /fwerte/, /axwera/ por /afuera/, /xarina/ por /arina/*
6. Arcaísmos léxicos del español estándar:  
     /*agora / <ahora>, /mesmo/, <mismo>, /andenantes/, <antes> /ansina/ <así>, /muncho/ <mucho> /ajga/ <haya>*
7. Diptongación de hiatos: \*  
     /*ljon/ <león>, /pjor/ <peor>, /twaya/ <toalla>, /kwete/ <cohete>*
8. Simplificación de grupos consonánticos:  
     /*dotor/ <doctor>, /eklise/ <eclipse>, /sétimo/ <séptimo>, /elétriko/ <eléctrico>*
9. Aféresis:  
     *pa' <para>, tá, <está>, ira <mira>, amá <mamá>*
10. Equivalencias acústicas:  
     /*gwéno/ <bueno>, /peksi/ <pepsi>, /abuxa/ <aguja>*
11. Metátesis:  
     /*swidad/ <ciudad>*
12. Generalización del artículo femenino con voces de origen griego terminadas en -a:  
     *la problema, la tema, la poema*

Estos rasgos presentes en el vernáculo de Los Ángeles se difundieron en la América española a partir de la colonia (siglos XVI-XVIII) y se han mantenido vivos en los pueblos y zonas rurales, alejados de las ciudades. En las últimas, a lo largo del tiempo, el español estuvo sujeto a cambios que no se extendieron en las zonas rurales (Lapesa 1981).

#### 4. CONCLUSIONES

En este trabajo he mostrado que resulta sumamente fructífero analizar el contacto de dialectos del español americano trasladados a los Estados Unidos tomando en cuenta el modelo tipológico de tierras altas y tierras bajas junto con el modelo sociolingüístico propuesto por William Labov (2001) durante la segunda mitad del siglo XX para recoger, analizar y explicar información pertinente al estudio de las hablas urbanas y sus

correlaciones sociales. Ello no solo ofrece continuidad entre los estudios dialectales tradicionales y los más innovadores sociolingüísticos, sino que permite explicar los cambios lingüísticos importantes debidos al contacto dialectal que ocurre en la población hispanohablante de Estados Unidos, específicamente en el condado de Los Ángeles, situado en el sur de California. En esta zona geográfica puede observarse que entre la población hispanohablante de clase trabajadora domina el español vernáculo de Los Ángeles, ya que los emigrantes de primera generación adoptan su vocabulario y algunos de ellos se vuelven bi-dialectales. Los hablantes de segunda generación suelen adquirir la variante vernácula de Los Ángeles, resultado de una koiné de dialectos del español mexicano que se conformó en los siglos XIX y XX. Ello sucede, independientemente del origen regional de los padres, siempre y cuando los hablantes de segunda generación estén en contacto con la comunidad chicana de clase trabajadora, la cual no sólo es mayoritaria, sino que goza de prestigio entre la población que habla español, también de clase trabajadora. El caso de los centroamericanos, en especial de los salvadoreños de clase trabajadora trasladados al condado de Los Ángeles, es contundente. La situación del español de Los Ángeles parece divergir del español de otros lugares de los Estados Unidos. Por ejemplo, en Chicago, a pesar de que predominan los hablantes de español mexicano, no se ha originado un vernáculo local de base mexicana en la clase trabajadora, según indica Potowski (2011), pero en Nueva York parece que se está formando un vernáculo (Otheguy y Zentella 2012).

Desde una perspectiva general, puede observarse que la adquisición de una variante dialectal, al igual que la de una lengua, es inmediata entre los hijos de extranjeros que se incorporan a una comunidad establecida. Los hablantes de segunda generación hablan de la misma manera que sus compañeros y amigos y no como sus padres (Labov 2001: 437 y ss.). Los hablantes de primera generación, en cambio, mantienen su dialecto original y, a veces, agregan a su habla rasgos del dialecto nuevo o suprimen elementos del propio dialecto, como resultado de la acomodación, que muchas veces es momentánea. El empleo de esta tipología de tierras altas y bajas permite distinguir los cambios que derivan del contacto dialectal de los cambios que surgen del contacto de lenguas, en este caso del inglés y el español, en los hablantes de español de los Estados Unidos.

## REFERENCES

- Aaron, Jessi y José Esteban Hernández (2007): "Quantitative evidence for contact induced accommodation: shifts in /s/ reduction patterns in Salvadoran Spanish in Houston", en Kim Potowski y Ricard Cameron (eds.): *Spanish in contact: policy, social and linguistic inquiries*. Amsterdam: Jon Benjamins, pp. 329-344.
- Allen, James y Eugene Turner (2002): *Changing faces, Changing places. Mapping Southern Californians*. Northridge: California State University, Center for Geographical Studies.
- Andrade, Argelia (2012): *Segmental and Intonational Evidence for a Los Angeles Chicano Spanish Vernacular*. UCLA, Ph.D. Dissertation.
- Ávila, Raúl (2009): "Los medios y el español: entre el inglés y las lenguas indoamericanas", en Airolti Colombo y Ma Ángeles Soler Arrechalde (coords.), *Normatividad y uso lingüístico*. México: UNAM, pp. 31-46.
- Bayley, Robert, Norma L. Cárdenas, Belinda Treviño Schouten y Carlos Martín Vélez Salas (2010): "Spanish Dialect Contact in San Antonio, Texas: An Exploratory Study", en Kimberly Geeslin y Manuel Díaz-Campos (eds.): *Selected Proceedings of the 14<sup>th</sup> Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 48-60.
- Bills, Garland y Neddy Vigil (2008): *The Spanish language of New Mexico and Southern Colorado. A linguistic atlas*. Albuquerque: University of New Mexico Press.
- Buzatu, Anamaria (2012): "Porteño Spanish meets Los Angeles Vernacular Spanish". Centro de Estudios del Español de Estados Unidos. UCLA. (ms)

- Canfield, Delos Lincoln (1981): *Spanish Pronunciation in the Americas*. Chicago and London: The University of Chicago Press.
- Carr, Jhonni (2013): "Linguistic Landscape: The Case of Huntington Park." *Why Spanish Matters III: 50 Shades of Spanish*. University of California, Los Angeles, May 16 2013. Conference presentation.
- Chambers, J.K (1992): *Sociolinguistic Theory*. Cambridge, MA: Blackwell.
- Chomsky, Noam (1986): *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger Publishers.
- Espinosa, Aurelio M (1911): *The Spanish Language in New Mexico and Southern Colorado*. Santa Fe, NM: New Printing Company.
- Fought, Carmen (2003): *Chicano English in Context*. New York: Palgrave Macmillan.
- Geoffroy Rivas, Pedro (1978): *La lengua salvadoreña*. San Salvador: Ministerio de Educación
- Giles, Howard y Nikolas Coupeland (1991): *Language: contexts and consequences*. California: Brooks/Cole Publishing.
- Henríquez Ureña, Pedro (1921). "Observaciones sobre el español en América", *Revista de Filología Española VIII*, pp. 357-390.
- Hernández, José Esteban (2002): "Accommodation in a Dialect Contact Situation", *Filología y Lingüística XXVIII (2)*, pp. 93-110.
- (2009): "Measuring rates of word-final nasal velarization: The effect of dialect contact on in-group and out-group exchanges", *Journal of Sociolinguistics 13/5*, pp. 583-612.  
<http://dx.doi.org/10.1111/j.1467-9841.2009.00428.x>
- y Rubén Armando Maldonado (2012): "Reducción de /s/ final de sílaba entre transmigrantes salvadoreños en el sur de Texas". *Lengua y migración 4, 2*, pp. 43-67.
- Kerswill, Paul (1996): "Children, adolescents, and language change". *Language Variation and Change 8*, pp. 177-202.  
<http://dx.doi.org/10.1017/S0954394500001137>
- (2000): "Koineization and Accommodation", en J.K. Chambers, Peter Trudgill y Natalie Schilling-Estes (eds.): *The Handbook of Language Variation and Change*. Malden, MA: Blackwell.
- Labov, William (2001): *Principles of Linguistic change*. Volume II: Social Factors. Oxford: Blackwell.
- Lamar Prieto, Covadonga (2012): *El español de California en el siglo XIX*. UCLA, Ph.D. Dissertation.
- Lapesa, Rafael (1981): *Historia de la lengua española*. Madrid: Gredos.

- Lipski, John (1994): *Latin American Spanish*. London and New York: Longman.
- (2008): *Varieties of Spanish in the United States*. Washington, D.C.: Georgetown University Press.
- Lope Blanch, Juan M. (2004): “Flujo y reflujo en el español de los Estados Unidos: Un caso particular”, *Cuestiones de filología mexicana*. México: UNAM, pp. 153-158.
- Méndez Pidal, Ramón (1962): “Sevilla frente a Madrid. Algunas precisiones sobre el español de América”, *Miscelánea homenaje a André Martinet*. Estructuralismo e historia.vol.3. Canarias: Universidad de La Laguna, pp. 99-165.
- Moreno de Alba, José (1994): *La pronunciación del español en México*. México: El Colegio de México.
- (2001): *El español en América*. México: Fondo de Cultura Económica.
- Otheguy, Ricardo, Ana Celia Zentella y David Livert (2007): “Language and Dialect Contact in Spanish in New York: Toward the Formation of a Speech Community”. *Language* 83 (4), pp. 770-802.  
<http://dx.doi.org/10.1353/lan.2008.0019>
- y Ana Celia Zentella (2012): *Spanish in New York. Language contact, dialect leveling and structural continuity*. Oxford: Oxford University Press.  
<http://dx.doi.org/10.1093/acprof:oso/9780199737406.001.0001>
- Páez Urdaneta, Iraset (1981): *Historia y geografía hispanoamericana del voseo*. Caracas: Casa de Bello.
- Parodi, Claudia (2004): “Contacto de dialectos en Los Ángeles: español chicano y español mexicano”, *Séptimo Encuentro Internacional de Lingüística en el Noroeste*. T 2. Hermosillo: Universidad de Sonora, pp. 277-293.
- (2006): “Multilingüismo y diglosia” *El español de América. Diatopía, diacronía e historiografía*. México: UNAM, pp. 433-456.
- (2008): “Stigmatized Spanish inside the classroom and out: A model of language teaching to heritage speakers”, *Heritage Language Education. A New Field Emerging*. New York and London: Routledge, pp. 199-214.
- (2009): “Normatividad y diglosia en Los Ángeles: Un modelo de contacto lingüístico”, en Airoldi Colombo y Ma Ángeles Soler Arrechalde (coords.), *Normatividad y uso lingüístico*. México: UNAM, pp. 47-67.

- (2011): “El otro México: español chicano, koineización y diglosia en Los Ángeles, California” *Realismo en el análisis de corpus orales*. Pedro Martín Butragueño, ed. México: El Colegio de México, pp. 217-243.
- (2014): “El español y las lenguas indígenas de los mexicanos en los Estados Unidos” *Historia de la sociolingüística en México*. Rebeca Barriga y Pedro Martín Butragueño, eds. México: El Colegio de México.
- Potowski, Kim (2008): “I was raised talking like my mom’: the influence of mothers in the development of MexiRicans, phonological and lexical features”. Jason Rothman and Mercedes Niño-Murcia (eds): *Linguistic identity and bilingualism in different Hispanic contexts*. Amsterdam: John Benjamins, pp. 201-220.
- (2011): “Intrafamilial dialect contact”, en Manuel Díaz-Campos (ed.) *The Handbook of Hispanic Sociolinguistics*. Sussex: Wiley-Blackwell, pp. 578-597.
- Quesada Pacheco, Miguel Ángel (2000): “El español de América Central”, en Manuel Alvar (ed.): *Manual de dialectología. El español de América*. Barcelona: Ariel, pp. 101-115.
- Raymond, Chase (2012): “Generational Divisions: Dialect Divergence in a Los Angeles-Salvadoran Household”, *Hispanic Research Journal* 13, 4, pp. 297-316.  
<http://dx.doi.org/10.1179/1468273712Z.00000000017>
- 2012a. “Reallocation of pronouns through contact: In-the-moment identity construction amongst Southern California Salvadorans”, *Journal of Sociolinguistics* 16/5, pp. 669-690.  
<http://dx.doi.org/10.1111/josl.12001>
- Romero, Matías (2003): *Diccionario de salvadoreñismos*. La Libertad, El Salvador: Editorial Delgado.
- Rosenblat, Ángel (1967): “Contactos interlingüísticos en el mundo hispánico: el español y las lenguas indígenas de América”, *Actas del II Congreso Internacional de Hispanistas*. Nimega, pp. 109-154.
- Sánchez, Rosaura (1994): *Chicano Discourse. Socio-historic Perspectives*. Houston: Arte Público Press.
- Silva-Corvalán, Carmen (1986): “Bilingualism and language change. The extension of *estar* in Los Angeles Spanish”, *Language*, 62, pp. 587-608.  
<http://dx.doi.org/10.1353/lan.1986.0023>  
<http://dx.doi.org/10.2307/415479>



- (1994): *Language contact and change*. Oxford: Clarendon Press.
  - y Manuel Gutierrez (1995): “On transfer and simplification: Verbal clitics in Mexican American Spanish”, en Peggy Hashemipour et al. (eds.): *Studies on Language learning and Spanish Linguistics in Honor of Tracy Terrell*. San Francisco: Mc Graw Hill, pp. 302-12.
- Veltman, Calvin (1983): *Language Shift in the United States*. Berlin: Mouton.  
<http://dx.doi.org/10.1515/9783110824001>
- (1990): The status of the Spanish language in the US at the beginning of the 21st Century. *International Migration Review*. 24, pp. 108-123  
<http://dx.doi.org/10.2307/2546674>
- Villa, Daniel J (2012): “Pos aquí nomás mis chicharrones truenan: El estatus del español en el sur de Nuevo México”, en Alicia Vargas y Ángeles Gallegos (eds.): *Culturas en movimiento*. Guadalajara: Universidad de Guadalajara, pp. 60-78.
- Woods, Michael y Susana Rivera Mills (2012): “El tú como un ‘mask’: Voseo and Salvadoran and Honduran identity in the United States”. *Studies in Hispanic and Lusophone linguistics* 5, pp. 191-219.
- Zentella, Ana Celia (1990): “Lexical leveling in four New York City Spanish dialects: linguistic and social factors”. *Hispania* 73, pp. 1094-1105.  
<http://dx.doi.org/10.2307/344311>
- (1998): *Growing up bilingual*. Oxford: Blackwell.





## Apéndice 1:

### CUESTIONARIO EMPLEADO PARA LAS ENCUESTAS

1. Tu nombre es...
2. ¿Qué edad tienes, en que fecha naciste?
3. ¿De dónde eres?
4. ¿A qué edad y en que fecha llegaste a los Estados Unidos?
5. ¿A qué edad y en que fecha llegaste a Los Ángeles?
6. Tu dirección es...
7. El barrio es...
8. ¿Cómo se llama el barrio, estado y país donde estaba tu casa cuando eras niño?
9. ¿Estas casado/a?
10. ¿Dónde nació tu padre?
11. ¿Vivió tu papá siempre allí o se cambió de lugar?
12. ¿Cuándo llegó tu papá a Los Ángeles?
13. ¿A qué se dedica tu papá? ¿qué estudió?
14. ¿Dónde nació tu mamá?
15. ¿Vivió tu mamá siempre allí o se cambió de lugar?
16. ¿Cuándo llegó tu mamá a Los Ángeles?
17. ¿A qué se dedica tu mamá? ¿Qué estudió?
18. ¿Qué idiomas hablabas de niño/a con tus padres?
19. ¿Qué idiomas hablabas de niño/a con tus hermanos?
20. ¿Qué idiomas hablabas de niño/a con tus amigos?
21. ¿Qué idiomas hablabas de niño/a en la escuela?
22. ¿Cuál es el lugar de origen de tu esposo/a, novio/a?
23. ¿Qué estudias? ¿Qué estudiaste?
24. ¿En qué trabajas?

25. ¿A qué lugares has viajado y por cuánto tiempo?
26. ¿Te gustaba el barrio donde vivías de niño? ¿por qué?
27. ¿Te gustaba el barrio donde vivías de adolescente? ¿por qué?
28. ¿Sabes alguna historia o anécdota de tu papá cuando era niño?
29. ¿Sabes alguna historia o anécdota de tu mamá cuando era niña?
30. ¿Sabes alguna historia o anécdota de tu abuelo cuando era niño?
31. ¿Sabes alguna historia o anécdota de tu abuela cuando era niña?
32. ¿Vas seguido a visitar a tus familiares a Latinoamérica u a otra ciudad de los Estados Unidos?
33. ¿Vienen a visitarte tu familiares de Latinoamérica o de otro lugar?  
¿De dónde vienen?
34. ¿Has estado tú allí?
35. ¿Sabes una historia de tus papás cuando vivían en Latinoamérica?
36. ¿Tienes algún recuerdo de tus amigos de cuando eras niño/a?
37. ¿Tienes algún recuerdo reciente de tus amigos?
38. ¿Qué idioma/s hablas con tu novio/a?
39. ¿Conoces el lugar de origen de tus papás? ¿Te gusta? ¿Por qué?
40. ¿Qué opinas de la educación en los Estados Unidos?
41. ¿Qué opinas de la secundaria (high school) en Los Ángeles?
42. ¿Dónde estudiaste la primaria y la secundaria?
43. ¿Qué opinas del español que se habla en Los Ángeles?
44. ¿Qué opinas de los mexicanos que viven en Los Ángeles? ¿Qué opinas de cómo hablan?
45. ¿Qué opinas de los salvadoreños que viven en Los Ángeles? ¿Qué opinas de cómo hablan?
46. ¿Qué opinas de los guatemaltecos que viven en Los Ángeles? ¿Qué opinas de cómo hablan?
47. ¿Qué grupos de latinos que viven en Los Ángeles te caen bien? ¿Por qué? ¿Qué opinas de cómo hablan?
48. ¿Qué grupos de latinos que viven en Los Ángeles te caen mal? ¿Por qué? ¿Qué opinas de cómo hablan?
49. ¿Qué opinas del voseo?
50. ¿Te gusta que te hablen de “vos”?

## Apéndice 2:

### REGIONALISMOS SALVADOREÑOS: ¿QUÉ QUIERE DECIR?

Preguntar la palabra en cursivas. Escribir sí cuando los informantes saben la respuesta y no, cuando no la saben.

Pregunta	Respuesta	Sí	No
1. <i>matate</i>	morral, bolsa		
2. <i>chuco</i>	sucio		
3. <i>guaro</i>	licor		
4. <i>piscucha</i>	papalote, cometa		
5. <i>marañero</i>	tramposo		
6. <i>ginas</i>	sandalias		
7. <i>ayote</i>	calabacita		
8. <i>cipote, bicho</i>	niño, muchacho, chamaco		
9. <i>socado</i>	apretado		
10. <i>fustán</i>	medio fondo		
11. <i>chumpa</i>	chamarra, saco		
12. <i>plátano</i>	plátano macho		
13. <i>guíneo</i>	plátano tabasco		
14. <i>majoncho</i>	plátano chiquito (dominico)		
15. <i>haragán</i>	perezoso		
16. <i>chabacán</i>	vulgar		
17. <i>chele</i>	rubio, güero		
18. <i>chambroso</i>	chismoso		
19. <i>pacha</i>	botella, mamila, biberón		
20. <i>keike</i>	pastel		

- |  |                                      |
|--|--------------------------------------|
| 21. <i>corvo, colín</i>                              | machete                              |
| 22. <i>cuchumbo</i>                                  | bote (de basura)                     |
| 23. <i>bayunco</i>                                   | que dice tonterías, bromista, simple |
| 24. <i>chacalines</i>                                | camaroncitos                         |
| 25. <i>tunco</i>                                     | cerdo                                |
| 26. ¿Usas estas palabras en Los Ángeles? ¿Con quién? |                                      |

## Remarks on Pronominal Perseveration and Functional Explanation

Ricardo Otheguy  
Graduate Center.  
City University of New York  
[rotheguy@gc.cuny.edu](mailto:rotheguy@gc.cuny.edu)

**Abstract:** Priming as the motivation for perseveration has been held up as an instance of mechanical, non-functional explanation in experimental and variationist linguistics. Perseveration due to priming has been claimed to account for, among others, features of morphosyntax such as the variable use of Spanish subject personal pronouns. The proposal is that the use in discourse of a pronoun with a finite verb token tends to lead to the subsequent use of another pronoun with the next finite verb token, and that the omission of a pronoun tends to lead to another omission. This paper makes two points, one theoretical, one empirical. The theoretical point relates to the inadequacy of generative models as accounts of language use, and consequently, their inappropriateness for the study of variation, including Spanish pronominal variation. The empirical point is that perseveration in the use of Spanish subject personal pronouns cannot be replicated in a new corpus to the degree that one might wish. While the omission of pronouns shows evidence of perseveration, their deployment does not. And perseveration in omissions, it is argued, does not require the postulation of structural priming; it can be accounted for by well-established, functional understandings of what subject pronouns do in discourse. Furthermore, even when structural priming is invoked, it does not qualify as a counter-functional explanation, once the boundaries of functionalism are properly delimited.

**Keywords:** Perseveration, priming, explanation, generative grammar, functional grammar, Spanish subject pronouns.



## 1. INTRODUCTION

In the study of many sociolinguistic variables, researchers have noted that speakers appear to persist in the production of one of the variants, making the same choice repeatedly over stretches of discourse before changing to the other variant. This tendency is known as *structural perseveration*, due to *structural priming*, or to a form of priming known as spreading activation. (the term *structural*, which we dispense with henceforth, distinguishes these tendencies and explanations from *lexical* perseveration and priming, which we will not discuss here). In this view, the deployment of a structural variant in a particular occasion of use is said to prime deployment of the same, perseverating variant in the next occasion of use. Perseveration due to priming has been part of research into variability in many languages, including the Spanishes that for several decades have been the masterfully cultivated field of Carmen Silva-Corvalán, to whom this volume is dedicated. The discovery of perseveration has led to important empirical and theoretical developments in variationist sociolinguistics. Yet there remain problems to be solved and misunderstandings to be cleared up. This note aims: (a) to examine the role of perseveration in Spanish subject personal pronouns in a new data set; (b) to use the results to assess critically a number of empirical generalizations that have been made regarding perseveration; and (c) to clarify the theoretical role of perseveration as a contributor to functional explanation in variationist sociolinguistics.

In Spanish phonology, perseveration has been argued to be a predictor of, for example, the choice between the presence of coda /-s/ (realized as [-s] or [-h]) and its absence (often indicated by a zero [-o]). Under perseveration, [-s] (or [-h]) is said to lead to more [-s] (or [-h]), whereas [-o] is said to lead to more [-o] (Poplack 1980, 1981, Terrell 1981). To use a standard illustration, in phrases like *las casas bonitas* ‘the pretty houses’, the perseverating pattern containing all [-s] (*la[s] casa[s] bonita[s]*), or containing all [-o] (*la[o] casa[o] bonita[o]*), is said to be favored over what we may call the *interspersing* pattern (*la[s] casa[o] bonita[o]*, or *la[o] casas[s] bonita[o]*).

In Spanish morphosyntax, evidence has been offered for regarding perseveration as a predictor of the choice between presence of a subject personal pronoun (e.g., *ella canta* ‘she sings’) and its absence (e.g., *canta* ‘she sings’) (see, among others, Cameron 1994, Cameron and Flores-Ferrán 2004). Perseveration is said to occur when the deployment of a pronoun in the first verb of a sequence leads to relatively more deployment in the next



verb (e.g., *ella canta pero yo no bailo* ‘she sings but I don’t dance’), and when the omission of a pronoun in the first verb leads to relatively more omission in the next (e.g., *canta pero no bailo*). The opposite of perseveration, which I call *interspersion*, is illustrated by deployment followed by omission (e.g., *ella canta pero no bailo*) and by omission followed by deployment (e.g., *canta pero yo no bailo*). While in principle perseveration due to priming can cover sequences of two, three, or more verbs, most research deals only with sequences of two, a practice that I will continue here. After the theoretical discussion in the next section, it is perseveration in deployed or omitted subject personal pronouns in sequences of two verbs in Spanish that is analyzed in the present work.

## 2. THEORETICAL BACKGROUND

The theoretical importance of perseveration in Spanishes derives from the fact that findings related to perseveration in coda /-s/ and subject personal pronouns have been used as part of the argument *against* the view that the internal predictor variables of sociolinguistic analyses have predominantly functional explanations. In two well-known and widely cited papers by the founder of the field (Labov 1987a, 1987b), the proposal is put forth, on the basis of findings by Poplack and Terrell, that perseverating Spanish speakers deploy and omit coda /-s/, not with an eye to sustaining meaning, a functional explanation, but rather in a fashion that Labov calls mechanical, a counter-functional explanation. In more general terms, perseveration is seen in a number of frequently cited works as a clear manifestation of functionally unmotivated linguistic behavior (Bock 1986, Bock and Griffin 2000, Weiner and Labov 1983).

Perseveration in deployment and omission of subject personal pronouns is thus an integral part of the thinking that sees choices between variants in mechanical, non-intentional, non-functional terms. Even though the seminal work on perseveration in Spanish subject personal pronouns by Cameron (1994) recognizes some functionality, Cameron sees perseveration as a mostly automated, non-functional process. In Cameron and Flores-Ferrán (2004) and Flores-Ferrán (2002), speakers are seen as deploying or omitting pronouns (in their terms, using an overt or a null pronoun) under the impetus of priming or spreading activation, which is explicitly regarded as an automatic process that leaves little or no room for intentionality. Even

authors of more recent works, who mainly study motivated uses of the pronouns, still assume that pronominal perseveration is a manifestation of non-functionality (Travis and Torres Cacoullos 2012: 733). From the earliest studies of perseveration in the variationist field, then, the tendency has been to background communicative intent and to highlight the automatic and the mechanical in the insertion or omission of pronouns.

But the counter-functional conclusions regarding perseveration by the authors cited above rest on an over-narrow definition of functionalism, making these conclusions more terminological than substantive. Under their definition, linguistic behaviors, such as choosing between [-s] and [-o], or between deployment and omission of a subject personal pronoun, are regarded as explainable functionally *only* when they can be connected to meaning or communication. In Labov (1987ab), for example, no other explanation is functional; the feature is either grounded on communicative need or is pushed over to the other, mechanical side, where it becomes part of a counter-functional account. This approach to the problem is of limited usefulness because functional explanations, while including communication, encompass much more. The point requires some elaboration.

The topic of whether or not a particular factor that statistically conditions the choice of variants is susceptible to functional explanation is most usefully raised in the context of ontological concerns about what language is and the kinds of explanations it requires. When these are the matters at issue, explanations are not simply either functional (because communicative) or mechanical. Rather, the discussion about functional and mechanical explanations enters into deeper questions regarding which features of language are involved in processes that are applicable to multiple areas of cognition, and which features are, in contrast, uniquely linguistic. The former explanations are *functional*; the latter explanations, the language-unique ones, are, let us call them, *foundational*. Functional explanations stem from a conception of language as a semiotic cognitive system that shares with other human systems large numbers of traits and constraining factors; foundational explanations stem from a conception of language as an autonomous computational system whose traits and constraining factors are unique to language. The difference between functional and foundational explanations, then, does not rest simply on whether appeals are made to communicative advantage. The difference, rather, is that functional explanations trade in factors external to language, while foundational explanations restrict their range to elements that are internal to it.

This understanding of functional explanation as encompassing much more than simply communicative need but covering rather all explanatory appeals to language-external considerations is widely shared (cf. Lindblom 1983: 218; Noonan 1998: 16). In this line of thinking, functional explanations in phonology and morphosyntax point to well-known physiological traits such as the relative size and mobility of different articulators, the shape of different cavities of the vocal tract, and the characteristics of the auditory apparatus. Functional explanations appeal as well to cognitive and psychological traits such as, to be sure, the need to communicate, but also to such factors as memory limitations and processing speed, as well as to preferences for, and tendencies toward, economy of effort, iconicity, egocentricity, and the affirmation of identity.

The defining trait of the functional explanation, then, is its formulation in terms of constructs whose applicability is not limited to language (so for example, the vocal tract is used for many other functions besides speech, while tendencies toward least effort and egocentricity are found in many domains other than language). In contrast to these many types of external functional explanation, foundational explanations are internal, in the sense of being articulated through terms whose applicability is unique to language, remitting to the linguistic genetic endowment and the acquisition biases resulting from the strictures of universal grammar.

Two examples help to make clear the distinction between functional and foundational explanations. The widely observed pattern of variation in many languages wherein coda obstruents are retained more frequently when the following word begins with a vowel than when it begins with a consonant is susceptible to a functional explanation. The phenomenon responds in large measure to the universal preference for syllables of consonant-vowel form (CV) over syllables consisting of consonant clusters. This preference has been explained, for speakers, in terms of synergetic motor constraints between different articulators and, in hearers, in terms of accommodation to perceptual and informational processing biases (Barry and Andreeva 2001: 57; Lindblom 1983: 232ff., 243). This explanation is functional because tendencies toward motor synergy and aural accommodation exist independently of language, and because the terms engaged in the explanation ('gestural synergy,' 'auditory biases') are not uniquely linguistic. In contrast, consider the generativist finding that rules are structure dependent (Chomsky 1971), that is, that they are not expressible in terms of words or morphemes, but must be stated rather in terms of categories such as NP and VP. This is a

foundational explanation because it is held to connect directly to properties of universal grammar with roots in the human genetic endowment, and because it is expressed in terms such as 'linguistic rule', 'NP category' and 'VP category' that are exclusively linguistic in character.

Little theoretical progress is to be made, then, by limiting functional explanations to those grounded in communication. Much greater profit is to be found in fixing the boundaries of functional explanation so as to encompass all accounts that are not directly derived from the arbitrary register of uniquely linguistic traits said to be found in the innate mental linguistic endowment. In this more encompassing view of functionalism, as soon as the explanation is grounded on attributes that are external to language, and as soon as it is formulated in terms of conditioning variables that have an impact on other human systems in addition to language, we have entered the realm of explanatory functionalism. The resident notions in this functional realm certainly include the interests of communication and meaning-making, but are by no means limited to them.

The foregoing suggests that we can best organize our thinking regarding these matters not only by recognizing the existence of two kinds of explanations (the functional and the foundational), but by recognizing also that the functional realm can itself be divided into two types, one grounded on communication, the other one on the vast array of physiological and psychological factors that impact on language. The distinction is implicit in all work on functionalism, and has been made explicit by several scholars, but perhaps most eloquently by Diver (1975). Taking an essentially functional approach in the broad sense advocated here, Diver urged on linguistics two distinct explanatory orientations, simply and directly called the *communicative factor* and the *human factor*. In these terms, while the explanations that we have called foundational are univocal, connecting in all cases to universal grammar, functional explanations are bivocal, appealing in some cases to the communicative factor and in others to the human factor. Since the need for effective communication is not the only functional factor sculpting the form of language, no counter-functional argument is offered by pointing to the lack of communicative advantage associated with a feature; many other types of advantages can still be recognized inside the boundaries of functionalism.

The present study of perseveration in Spanish subject personal pronouns, then, is meant as a contribution to the dialogue regarding explanation in linguistics, a contribution that will center around a number

of empirical clarifications regarding pronominal perseveration. The study confirms some previous findings, qualifies others, and attempts to clarify what these findings tell us about the realm of explanation to which appeal must be made in order to understand Spanish pronominal use.

### 3. PERSEVERATION IN SPANISH SUBJECT PERSONAL PRONOUNS: DATA SELECTION AND METHOD OF ANALYSIS

Three preliminary clarifications are in order. First, the general practice, adopted here, of studying subject personal pronouns, including their perseveration, on the basis of the alternation in finite verb tokens between presence and absence of the pronoun (e.g. *ella canta* ~ *canta*) must be acknowledged to introduce several notable simplifications. The approach leaves out of consideration those subject personal pronouns that in many Spanishes occur with infinitives (e.g., *para ella cantar mejor tendría que practicar más* ‘in order for her to sing better she would have to practice more’). The approach also excludes finite verb tokens occurring with nominal subjects (*María canta* ‘Maria sings’). And it counts the over 90 to 95 percent of finite verb tokens occurring preposed to the verb (*ella canta*) together with the small percentage occurring postposed (*canta ella*); or, alternatively, leaves the latter out altogether. All this is simply to say that, in most studies, the linguistic variable is taken to have two variants, pronominal presence versus pronominal absence (often expressed in terms of overt pronoun versus null pronoun). While this two-variant approach leaves room for debate, I find it useful enough for the purposes of most studies (including this one) and will adopt it here without further justification.

Second, Spanish subject personal pronouns are used in corpora, in proportions of 90 percent or higher, for the purpose of animate reference, making it useful to proceed, as do most studies, including the present one, as if they were completely restricted to such references. Third, in most, perhaps all, corpora verb tokens appearing with a pronoun are in a minority, most verbs occurring with only an inflectional indication of the subject (this is so even in high-pronoun Spanishes such as those in the Dominican Republic). Since the present study does not make a deep-surface distinction, it can properly be said, and will be important to insist on saying for the arguments that follow, that the vast majority of Spanish verb tokens, the inflectional suffix aside, occur *without* a subject (that is, not with a null pronoun subject).

We represent verb tokens like *canta* as [V] and verb tokens like *ella canta* (and the small percentage of *canta ella*) as [pV]. Nominal subjects, including infinitival and clausal ones, are represented as [nV]. Figure 1 represents sequences that instantiate perseverance and sequences that do not. I call the latter *interspersion*, to capture the sense of an insertion of a mismatched structural element that interrupts what could have been a matching, structurally perseverating sequence. Perseveration, regularly attributed to priming, is said to be demonstrated when the frequency of matching or perseverating sequences is relatively higher than that of unmatching or interspersing sequences.

Sequences of *perseveration*: [V – V], [pV – pV]

Sequences of *interspersion*: [pV – V], [V – pV], [nV – V], [nV – pV].

**Figure 1.** Finite verb sequences instantiating pronominal perseverance and interspersion.

The literature often refers to the first verb of a sequence, the one doing the priming, as the *trigger*, and to the second verb, the one being primed, as the *target* (Cameron 1994). But these terms, as we shall see, introduce implicit analyses that will be questioned here. I prefer to speak more plainly of the first and second verbs of a sequence (V1 and V2 for short). The literature also tends to refer to cases like *ella canta* as having an *overt* pronoun, and to cases of *canta* as having a *null* pronoun, often indicated with a zero. These terms too are theory-dependent and introduce implicit analyses. As I have been doing above, I prefer to continue to speak, in simpler and more observational terms, of verbs where the pronoun is present, used, deployed, etc., and of verbs where the pronoun is absent, omitted, not deployed, etc. (Otheguy and Zentella 2012:9n).

Spanish pronominal perseverance is studied here using data drawn from the Otheguy-Zentella corpus (henceforth OZC), gathered in New York City (Otheguy and Zentella 2012). From the OZC, eight interviews were selected, extracting from each the first 100 verb tokens (for a total of 800) that represent sites of pronominal variation [pV] ~ [V]. We refer to these as *eligible* verb tokens, to distinguish them from ineligible tokens that do not constitute sites of pronominal variation. Examples of eligible tokens (and using only examples of preposed pronouns) are thus *ella canta* and *canta* ‘(she) sings’; *el artista que ella contrató canta* and *el artista que contrató canta* ‘the artist that (she) hired sings’; and *ella se mueve* and *se mueve* ‘(she) moves’.



In contrast, verb tokens that are ineligible for the study of pronominal variation, and cannot therefore be V2 in sequences analyzed for perseveration, are those belonging to types where there is little or no [pV] ~ [V] alternation. These include verb tokens already showing nominal subjects [nV], as in *María canta*, where it is rare to find *María ella canta*; or clausal subjects, as in *el artista que ella contrató canta*, where one seldom finds *el artista que ella contrató él canta*. Also ineligible are those tokens of [V] found in contexts where [pV] seldom occurs, such as subject-headed relative clauses, as in *el artista que aceptó el contrato canta* ‘the artist who accepted the contract sings’, where it is rare to find *el artista que él aceptó el contrato canta*; or found in clauses of any type making reference to the weather, as in *nieva mucho* ‘it snows a lot’, where *ello nieva mucho* is not frequent; or found, as mentioned above, in cases when the subject refers to inanimates, as in *se mueve*, when expressing ‘it moves’. (More details on the formation of the corpus and the difference between eligible and ineligible verbs are in Otheguy and Zentella 2012: Chapter 3).

In establishing whether a V2 (an eligible finite verb token that is second in a sequence) should be counted as an instance of perseveration or interspersion, my method departs somewhat from recent approaches, for reasons discussed presently. In order to determine whether there is a perseverating match or an interspersing mismatch, I compare each V2 (whether [pV] or [V]) with the *nearest preceding animate V1*, whether coreferential or not, eligible or not. In addition, and as long as it is animate and nearest, this V1 can, but does not have to, be uttered by the informant; it could be, and often is, uttered by the interviewer (thus covering both long narratives and back-and-forth dialogue). The V1, then, is the nearest animate coreferential or non-coreferential [V], or [pV], or [nV] used by the informant or the interviewer. This approach produces a majority of adjacent V1 – V2 sequences, but allows for cases where the V1 is several clauses back, due to intervening verb tokens whose subjects are inanimate.

In most recent work, perseveration or interspersion has been established a little differently, restricting V1 to coreferential, and sometimes to coreferential and eligible, verb tokens (Abreu 2012, Carvalho and Child 2011, Flores-Ferrán 2002, Travis 2007). This allows, in the more explicit statements of method, for up to two verb tokens with animate, non-coreferential subjects to intervene between V2 and the potentially priming V1 (Travis and Torres Cacoullós 2012: 730). In addition, some studies appear to require not only eligible V2’s, as I do, but also eligible V1’s, skipping over all [nV]’s as potential V1’s.

The reason I accept as V1 all tokens making animate reference, ignoring considerations of coreference, is that I am examining the general question of the purported mechanical and non-functional character of perseveration as a motivation for the choice of linguistic variant (Labov 1987ab). An analysis of structural perseveration that takes coreference into account, useful as it may be for other purposes, does not isolate a linguistic feature of clear relevance to the difference between functional and foundational explanations, for the simple reason that coreference concerns the relation between linguistic objects and features of the external world in acts of language use. Once coreference is brought into the picture, functional effects and matters of communicative efficacy start to play a role, as has been clearly recognized from the earliest studies of perseveration (Cameron 1994: 40, 1995: 10). Once this happens, perseveration ceases to represent the completely non-functional process that it is often taken to be and that provides it with theoretical interest.

In addition, the reason that I adhere to a strict requirement of nearness for V1 (as long as they are animate), ignoring eligibility of V1's and including the V1's of the interviewer, is that it is the influence of this nearest animate V1, coreferential or not, eligible or not, uttered by the informant or the interviewer, where structural perseveration as a purely mechanical tendency should logically manifest itself with the greatest force. That is, the reason that I take a mechanical approach to the identification of V1 is that I want to test whether priming is the mechanical conditioning factor that it is claimed to be, and to do so in what should be the most favorable environment to produce perseveration. Finally, the reason that, in my approach, [nV] and all other ineligible tokens that cannot be V2 can however be V1 is that the question being pursued is whether the form of the subject of a V2, either [pV] or [V], is in some way causally connected to the form of the preceding V1; for that purpose, all animate verb tokens are grist for the mill, and all should be considered V1, even if they are of a type that rules them out as V2. To be sure, all V2's preceded by a V1 of the form [nV] necessarily become cases of interspersion ([nV] – [pV] or [nV] – [V]). But they do so coherently, as in each case the form of the V2 is properly tested for the causal connection through priming with the form of the V1.

The texts (1) through (3) are drawn from an interview in the OZC with a Colombian newcomer included in our sample. They are used here to illustrate eligible and ineligible verb tokens and, within the eligible tokens, to illustrate perseveration and interspersion. The speaker's identification



number appears in brackets at the end of each of the texts (the letter after the number indicates the country of origin). The English translation follows that.

- (1a) [A discussion of what to do with found money] Yo siempre he oído que el que tiene dinero olvida todas las cosas que piensa, mas se lo digo, en lo que yo pienso, lo primero que yo haría sería ayudar la familia [021C]. ‘I’ve always heard that the person who has money forgets everything he thought, but I tell you what I think, the first thing I would do is help the family.’
- (1b) Si usted me pregunta, yo le puedo contestar algo más, ¿no? [021C]. ‘If you ask me I can answer something more, right?’
- (2) Porque el padre no le dedica el tiempo, el poco tiempo que le queda no se lo dedica. Lo utiliza para otras cosas, y entonces el muchacho se siente libre y como todo joven es curioso y busca, busca siempre probar una u otra cosa nueva. Pienso que ahí está la diferencia, en la forma como aquí trabajamos y en la forma como en nuestro país trabajamos [021C]. ‘Because the father does not devote the time, the little time left to him, (he) does not devote it. (He) uses it for other things, and then the boy feels free and like all young people is curious and (he) looks, (he) looks to try something new. (I) think that there is the difference, in the way (we) work here and the way (we) work in our country.’
- (3) Aquí en Estados Unidos en un tren yo noté que una persona, un señor se acercaba mucho a mí, me empujaba, me empujaba, cuando yo sentí el ruido de la cremallera en mi cartera que estaba corriendo, volteé a mirar y claro el señor estaba abriendo mi cartera. Así que entonces yo retiré mi cartera, la cerré, lo miré a él. Él se debió de dar cuenta de que lo había descubierto ... [021C]. ‘Here in the U.S. on a train I noticed that a person, a man was getting too close to me, (he) pushed me, (he) pushed me, when I heard the noise of the zipper in my wallet that was moving, (I) turned around to look and of course the man was opening my wallet. So I pulled back my purse, (I) closed it, (I) looked at him. He must have realized that (I) had caught him ...’

In these examples, note first the difference between eligible verb tokens (all underlined) and ineligible verb tokens (not underlined). Eligible verb tokens in (1a) include, among others, *Yo siempre he oído*,

which could've been *siempre he oído*; *piensa*, which could've occurred as *él piensa*; and *se lo digo*, which could've been *yo se lo digo*. The two verb tokens in (1b) are eligible, since the phrase *si usted me pregunta yo le puedo contestar*, where both verbs occur with pronouns, could have occurred without pronouns, *si me pregunta le puedo contestar*. Eligible tokens underlined in (2) include the second *dedica*, which could have been él *no se lo dedica*, and *Pienso*, which could have been Yo *pienso*. Examples of eligible tokens in (3) include *yo noté* and *me empujaba*, which could've been *noté* and *él me empujaba*.

In contrast, consider the verb tokens that are ineligible (and not underlined). In (1a), examples of ineligible verb tokens are *tiene*, whose subject is a headless relative; *olvida*, which already has a clausal subject; and *sería*, which does not have an animate subject. In (2) the first *dedica* is ineligible because it appears with a nominal subject; the same is true for *se siente*. In (3), examples of ineligible tokens include *se acercaba* and *estaba abriendo*, which have nominal subjects.

Note now examples of perseveration and interspersions. In (1a), cases of perseveration are [*piensa/digo*], and [*yo pienso/yo haría*]; cases of interspersions are [*el que tiene dinero olvida/piensa*] and [*digo/yo pienso*]. In (1b) there is only the one sequence, and it represents perseveration [*usted pregunta/yo puedo*]. In (2), cases of perseveration are [*dedica/utiliza*], [*busca/busca*], [*busca/Pienso*], [*Pienso/trabajamos*] and [*trabajamos/trabajamos*]; cases of interspersions are [*el poco tiempo que le queda/dedica*] and [*todo joven es/busca*]. In (3), cases of perseveration are [*empujaba/empujaba*], [*estaba corriendo/voltée a mirar*], [*cerré/miré*]; cases of interspersions are [*un señor se acercaba/empujaba*], [*empujaba/yo sentí*], [*yo sentí/estaba corriendo*], [*estaba abriendo/yo retiré*], [*yo retiré/cerré*], and [*miré/él debió*].

#### 4. PERSEVERATION IN THE DOMAIN OF SPANISH SUBJECT PERSONAL PRONOUNS: OVERALL QUANTIFICATION

The distribution of perseveration and interspersions is presented in Table 1, distinguishing between sequences where the second verb is [pV] (left side panel) and those where the second verb is [V] (right side panel). Across both panels, the first line shows cases of observed perseveration and the second one shows cases of observed interspersions.

	Second verb is [pV]		Second verb is [V]			
		N	%	N	%	
Perseveration	[pV pV]	102	41	[V V]	392	71
Interspersion	[V pV] or [nV pV]	146	59	[pV V] or [nV V]	160	29
		248	100		552	100
x <sup>2</sup> = 64.707		p < .0001				

**Table 1.** Perseveration and interspersion, by [pV] and [V]

The table shows the radical lack of equivalence between [pV] and [V]. The dominant observable pattern for [V] is perseveration whereas the dominant pattern for [pV] is interspersion. At this level of observation, the perseveration hypothesis could be said to find support in cases of [V], because the majority of preceding verbs (71 percent) are also [V]. But the hypothesis finds little support in cases of [pV], because the majority of preceding verbs (59 percent) are not [pV]. Perseveration does not appear from Table 1 as a general phenomenon, but one that is concentrated mostly on [V – V].

## 5. PERSEVERATION IN THE DOMAIN OF SPANISH SUBJECT PERSONAL PRONOUNS: THE REFERENCE VARIABLE

Studies of perseveration usually take into account whether the second verb of a potentially perseverating sequence is coreferential with the first. Generally speaking, and in one of the most consistent results of Spanish corpus-based studies, the proportion of verbs with [pV] is always *lower* in contexts of same reference than in contexts of switch reference (Bentivoglio 1987, Cameron 1994, 1995, Hochberg 1986, Otheguy and Zentella 2012, Silva-Corvalán 1982).

Note that coreference is distinct from perseveration. In sequences in (1) above we find them as follows: [*cerré/miré*] same reference, perseveration; [*yo retiré/ cerré*] same-reference, interspersion; [*empujaba/ yo sentí*] switch-reference, interspersion; [*estaba corriendo/ volteé a mirar*] switch reference, perseveration. In Cameron (1994), the perseveration effect was concentrated in cases of same reference. In subsequent studies, it is usually only these same reference cases that are analyzed.

### 5.1. Perseveration in same-reference contexts

In what follows, we separate perseveration results by contexts of same- and switch- reference. Results for switch-reference contexts are in the next section. Results for same-reference contexts are here (Table 2).

		Second verb is [pV]		Second verb is [V]		
		N	%		N	%
Perseveration	[pV pV]	64	49	[V V]	241	74
Interspersion	[V pV] or [nV pV]	68	51	[pV V] or [nV V]	84	26
		132	100		325	100
x <sup>2</sup> = 27.864		p < .0001				

**Table 2.** Perseveration and interspersion, by [pV] and [V] (same reference only)

Table 2 shows a pattern similar to that of Table 1. Simple observed perseveration obtains in contexts of same reference for instances of [V], but not for [pV]. Again here, in contexts of same reference, the perseveration hypothesis is not strongly confirmed; it registers a simple observed statistical favoring only for verbs used without pronouns, [V], but not for verbs used with pronouns, [pV]. This is a departure from the literature, since it is in these same-reference contexts where Cameron (1994) most clearly locates a perseverating effect.

### 5.2. Perseveration in switch-reference contexts

Now we look at perseveration in contexts of switch reference, where the pattern that we have been observing appears again (Table 3).

		Second verb is [pV]		Second verb is [V]		
		N	%		N	%
Perseveration	[pV pV]	36	32	[V V]	148	67
Interspersion	[V pV] or [nV pV]	76	68	[pV V] or [nV V]	72	33
		112	100		220	100
x <sup>2</sup> = 37.073		p < .0001				

**Table 3.** Perseveration and interspersion, by [pV] and [V] (switch reference only)

Here again perseveration appears concentrated in [V], where the perseveration rate is 67 percent. In contrast, the table shows that for [pV] the prevailing pattern is interspersion; the perseveration rate for [pV] is only 32 percent.

Tables 1, 2, and 3 offer interpretable results. On the basis of simple observations, the hypothesis, as a general expectation of perseveration for both [pV] and [V], does not appear to hold up well in any context. Throughout, the perseverating tendencies predicted by the hypothesis are observed for verbs without pronouns, [V], but not for verbs with pronouns, [pV]. For cases of [pV], perseveration is less frequent than, or at best as frequent as, interspersion. In contrast, for cases of [V], perseveration is the prevailing observable pattern. This conclusion gains further support when we study the eight informants individually. The informants that use more [V] tend to be the same as those in whom more perseveration is found. The top three users of [V] are, in the same order, the top three perseverators. Along the same lines, three of the four bottom users of [V] are the same as the bottom three perseverators. In short, there is no evidence, in this sample of the OZC, for a generalized priming effect on both [V] and [pV]. Moreover, as we argue presently, a priming interpretation would be misguided even for the perseverating [V]'s.

The conclusions drawn from Tables 1 – 3 are subject to some caveats. First, my coding for perseveration and coreference differs from some of the literature in that for me the V1 must be animate and nearest, and in that my work does not take into account reference chains and set-to-elements considerations, as in Cameron (1995). Though unlikely, this could have skewed my results in Table 2. Second, perseveration results here, as in the literature in general, are observational; a more sophisticated statistical analysis of the difference between expected and observed perseveration is for another time. Third, even though my results are all statistically significant, they may change when a larger proportion of the OZC is covered. Fourth, and finally, some of the literature relies on multivariate logistic regression run on Varbrul, whereas here we have only presented bivariate cross-tabulations. Run on background, the multivariate results that I obtain (run on SPSS) do not alter the picture offered here. I have chosen to show the bivariate results because it is this mode of presentation that best highlights the difference between [pV] and [V] that is relevant to my point, a difference that tends to be obscured in the output of regressions.

When the perseveration hypothesis is studied, as we have done here, under conditions that are most favorably inclined to confirm it, namely in

terms of the immediately preceding verb with an animate subject, the results suggest that the literature may be overstating the case. The big, often stated generalization that, at least in some contexts, the presence of a pronoun leads to more pronouns is not confirmed in this sample of the OZC.

## **6. PERSEVERATION IN THE DOMAIN OF SPANISH SUBJECT PERSONAL PRONOUNS: A FUNCTIONAL ACCOUNT**

What is the explanation for the simple fact that perseveration is observationally valid for [V] but not for [pV]? Why are the sequences of [V – V] the observed majority in all contexts, but not so the sequences of [pV – pV]? The answer seems straightforward. When they occur in running discourse, the person-number inflections of Spanish verbs provide sufficient information for the identification of subjects. The pronouns are thus best seen as something of a specialized extra tool that is used when something in the communicative task calls for greater than normal explicitness with regard to the question of what, or who, is the subject of the verb.

Note that there is no circularity here. I am *not* saying that we know that the pronoun is deployed because it is especially needed and that we know that it is especially needed because it is deployed. Rather, the general claim that pronoun use responds to communicative discourse pressures is supported by the fact that most of the statistical predictors of Spanish pronoun deployment are in fact straightforwardly related to reference identification in running discourse under clearly identifiable circumstances. For example, and as we have mentioned, pronominal deployment is always proportionately lower in contexts of same reference than in contexts of switch reference. Many other similar, functionally motivated distributions are documented in the literature (for details and references, see Otheguy, Zentella and Livert 2007; Otheguy and Zentella 2012).

The valid, non-circular generalization is that when special problems related to reference identification and tracking are not present, the default usage is the verb without a pronoun. It is thus only to be expected that a statistical check would identify a strong statistical perseveration effect of pronominal omission. The many perseverating sequences of verb tokens without pronoun that are found in Spanish corpora exist as simple manifestations of the fact that, in verb token after verb token, there is no reason to resort to the somewhat specialized tool represented by the pronoun.

Perseveration of pronominal omission is thus a statistical fact that is already well understood, and that is motivated by familiar communicative factors. As such, it hardly qualifies as convincing evidence of priming, or as evidence for the mechanical, and against the intentional, in theoretical discussions of pronominal use. Saying it another way, observed perseveration, concentrated as it is for the most part in [V], and representing only a minority of observable cases in [pV], does not warrant the appeal to priming as a cognitive mechanism, given that simpler and better-understood characterizations of the statistical facts are ready at hand.

In need of highlighting here are the two terminological matters mentioned above. As we have been arguing, the explanation for the prevailing perseveration of [V] is the simple fact of consistently dispensing with a linguistic tool, the subject pronoun, in the many repeated instances when it is not needed. Such tokens most decidedly do not occur *with* a 'null' pronoun. When the terms null and overt are used, the conceptualization is of a verb token that occurs with either of two allomorphs of the same unit, the always-present pronoun, in either its null or overt manifestation, a conceptualization that would render inoperable the account that we have just given of why sequences of [V – V] are so common. In the present proposal, the [V]'s are not verbs with a null pronoun, but rather verbs without a pronoun, or any other kind of subject, save the inflectional suffix.

Similarly, the terms 'trigger' and 'target' imply that the mere statistical fact of perseveration justifies the postulation of priming. The terms themselves assume the mechanical, even though, as we have seen, it is communicative-based choice that is much more likely to be involved. That is, the terms trigger and target assume that the form of V1 is automatically, and without intention, triggering the form of V2 when, much more likely, the speaker is deciding, for each V2, whether it needs to be [pV] or can be in the simpler form [V].

To the extent that, in some V2's, priming may be playing a role in the use of some matching tokens of [V] and, more plausibly perhaps, in some matching tokens of [pV] (which, to recall, are the statistical minority), it is well to recall that functional explanations are not limited to meaning and communication, but enlist any human trait that influences perceptions and behaviors outside the linguistic realm. It should be clear then that Spanish pronominal perseveration, to the extent that it exists in [pV], and insofar as it provides some evidence for priming, cannot be used as an argument against the prevalence of functional explanation in language. To the contrary,



perseveration would itself be very much susceptible to an external, functional account. For example, Cameron and Flores-Ferrán (2004), in explaining perseveration in terms of spreading activation theory (a form of priming), make reference to Bock (1986), who regards perseveration as a facilitative relationship in the interplay of production and comprehension; the speaker who perseverates in procedures that have been already activated may be reducing processing demands on the hearer.

More directly tied to the wider conception of functional explanation that we have insisted upon here is the work by Dell, Burger and Svec (1997), who connect their model of perseverative speech errors to non-linguistic areas of cognition, such as the playing of musical instruments (1997:141), and who explicitly regard perseveration as functional due to its relevance in accounts of memory and motor control (1997: 123). In the same vein, Travis and Torres Cacoullos (2012) cite van der Wel, Fleckenstein, Jax, and Rosenbaum (2007), who document the prevalence of perseveration in human motor activity. Far from supporting a turn in sociolinguistics away from the functional and toward the mechanical, and through the mechanical to the foundational, pronominal perseveration, to the extent that it exists, would appear as precisely the opposite, as an instantiation of the kind of multiply applicable human trait that enhances the scope of functionalism in the explanation of linguistic phenomena.

## **7. DISCUSSION AND CONCLUSION**

The foregoing has pointed out (a) the failure to replicate empirical findings related to Spanish finite verb tokens found in corpora as [V] and [pV]; in the sample of OZC studied here, observed perseveration is the prevailing pattern only for [V], registering in only a minority of cases in [pV], contrary to prior results; and (b) the misinterpretation of perseveration, such as it exists, as supporting a mechanical rather than a motivated view of variation, and as calling for a foundational rather than a functional approach to explanation.

These problems stem, to a considerable extent, from the mistaken belief that the theory underlying variationist sociolinguistics is a Chomskyan generative grammar, a computational system that is explicitly formulated to be indifferent to language use and communication, and that derives its strength and interest precisely from this indifference. The belief that sociolinguistics



connects to generativism was formulated in the earliest days of the field (Cedergren and Sankoff 1974) and is current to this day (Chambers 2009). The belief encourages the search for (and the predisposition to find) facts that seem automated and unrelated to communication, such as the Spanish sequences of [V – V] and [pV – pV]. It then encourages accounts of these pseudo facts in mechanical terms, such as the automatic drive to persevere under the auspices of priming. These explanations are consequently then seen as removed from the domain of functionalism, as attached to the foundational realm, and as thus linked to the generative enterprise.

But the chain of reasoning is flawed. The facts, as in the case discussed here, are often not what they appear to be, but hold true in much smaller magnitudes than proponents believe, or not at all. The problem in the present case is caused in part by the mechanistic terminology of ‘null’ and ‘overt’, and ‘trigger’ and ‘target’, which make it difficult to identify the interspersion that characterizes the use of [pV] and to understand the perseveration that characterizes the use of [V].

The priming explanation, even in the small number of cases where it may be applicable, is, in any case, itself not foundational, but functional. As such, it relates badly, if at all, to the principles of generative theory. For the study of perseveration / interspersion, as for the study of use in general, including both variable and categorical uses, a different theoretical infrastructure is needed. Such an infrastructure would have to be drawn from the many found in the theoretical marketplace of today, where theories abound that draw their facts from usage and allow for their constructs to connect more naturally with functional explanations.

## REFERENCES

- Abreu, Laurel (2012): "Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish", in Kimberley Geeslin and Manuel Díaz-Campos (eds.): *Selected Proceedings of the 14<sup>th</sup> Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 1-8.
- Barry, William and Bistra Andreeva (2001): "Cross-linguistic differences in spontaneous speech patterns", *Journal of the International Phonetic Association* 31, pp. 51-66.  
<http://dx.doi.org/10.1017/S0025100301001050>
- Bentivoglio, Paola (1987): *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela.
- Bock, J. Kathryn (1986): "Syntactic persistence in language production". *Cognitive Psychology* 18, pp. 355-387.  
[http://dx.doi.org/10.1016/0010-0285\(86\)90004-6](http://dx.doi.org/10.1016/0010-0285(86)90004-6)
- and Zenzi Griffin (2000): "The persistence of structural priming: Transient activation or implicit learning". *Journal of Experimental Psychology: General* 129, pp.177-192.  
<http://dx.doi.org/10.1037/0096-3445.129.2.177>
- Cameron, Richard (1992): *Pronominal and null subject variation in Spanish: Constraints, dialects and functional compensation*. University of Pennsylvania Ph.D. dissertation.
- (1994): "Switch reference, verb class, and priming in a variable syntax", in K. Beals et al (eds.): *Papers from the 30th Regional Meeting of the Chicago Linguistic Society: The Parasession on Variation in Linguistic Theory*. Chicago Linguistic Society 2, pp. 27-45.
- (1995): "The scope and limits of switch reference as a constraint on pronominal subject expression", *Hispanic Linguistics* 6, pp. 1-27.

- and Nydia Flores-Ferrán (2004): “Preservation of subject expression across regional dialects of Spanish”, *Spanish in Context* 1, pp. 41-65.  
<http://dx.doi.org/10.1075/sic.1.1.05cam>
- Carvalho, Ana M. and Michael Child (2011): “Subject pronoun expression in a variety of Spanish in contact with Portuguese”, in . In Jim Michnowicz and Robin Dodsworth (eds.): *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Press, pp. 14-25.
- Cedergren, Henrietta and David Sankoff (1974): “Variable rules: Performance as a statistical reflection of competence”, *Language* 50, pp. 233-255.
- Chambers, J. K. (2009): *Sociolinguistic theory* (Revised edition). Wiley-Blackwell Publishers.
- Chomsky, Noam (1971): *Problems of knowledge and freedom*. London: Pantheon Publishers.
- Dell, Gary, Lisa Burger, and William Svec (1997): “Language production and serial order: A functional analysis and a model”, *Psychological Review* 104, pp. 123-147.  
<http://dx.doi.org/10.1037/0033-295X.104.1.123>
- Diver, William (1975 [2012]): “The nature of linguistic meaning. *Language: communication and human behavior: The linguistic essays of William Diver*”, in Alan Huffman and Joseph Davis (eds.): Boston: Brill Publishers, pp. 47-64.
- Flores-Ferrán, Nydia (2002): *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City*. Munich: Lincom Europa.
- Hochberg, Judith (1986): “Functional compensation for /s/ deletion in Puerto Rican Spanish”. *Language* 62, pp. 609-621.  
<http://dx.doi.org/10.2307/415480>  
<http://dx.doi.org/10.1353/lan.1986.0041>
- Labov, William (1972): *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
  - (1987a. [1994a]): “The overestimation of functionalism”, *Principles of linguistic change: Internal factors*. Oxford: Blackwell Publishers, Chapter 19.
  - (1987b. [1994b]): “The maintenance of meaning”, *Principles of linguistic change: Internal factors*. Oxford: Blackwell Publishers, Chapter 20.

- Lindblom, Björn (1983): “Economy of speech gestures”, in Peter MacNeilage (ed.): *The production of speech*. New York: Springer-Verlag, pp. 217-245.  
[http://dx.doi.org/10.1007/978-1-4613-8202-7\\_10](http://dx.doi.org/10.1007/978-1-4613-8202-7_10)
- Noonan, Michael (1998): “Non-structuralist syntax”, in Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan and Kathleen Wheatley (eds.): *Functionalism and formalism in linguistics*. Philadelphia: John Benjamin Publishing Co., pp. 11-32.
- Ohala, John J. (1983): “The origin of sound patterns in vocal tract constraints”, in Peter MacNeilage (ed.): *The production of speech*. New York: Springer-Verlag, pp. 189 - 216.  
[http://dx.doi.org/10.1007/978-1-4613-8202-7\\_9](http://dx.doi.org/10.1007/978-1-4613-8202-7_9)
- Otheguy, Ricardo and Ana Celia Zentella (2012): *Spanish in New York: Language contact, dialectal leveling and structural continuity*. Oxford University Press.
- Ana Celia Zentella and David Livert (2007): “Language and dialect contact in Spanish in New York: Toward the formation of a speech community”, *Language* 83, pp.770-802.  
<http://dx.doi.org/10.1353/lan.2008.0019>
- Poplack, Shana (1980): “The notion of plural in Puerto Rican Spanish: Competing constraints on /s/ deletion”, in William Labov (ed.): *Locating language in time and space*. New York: Academic Press, pp. 55-68.
- (1981): “Mortal phonemes as plural morphemes”, in David Sankoff and Henrietta Cedergren (eds.): *Variation omnibus*. Edmonton, Alberta: Linguistic Research, pp.59-72.
- Shin, Naomi Lapidus and Ricardo Otheguy (2013): “Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York”, *Language in Society* 42, pp. 429-452.  
<http://dx.doi.org/10.1017/S0047404513000468>
- Silva-Corvalán, Carmen (1982): “Subject expression and placement in Mexican-American Spanish”, in Jon Amastae and Lucía Elías-Olivares (eds.): *Spanish in the United States: Sociolinguistic aspects*. Cambridge University Press, pp. 93-120.
- Terrell, Tracy (1981): “Diachronic reconstruction by dialect comparison of variable constraints: -s aspiration and deletion in Spanish”, in David Sankoff and Henrietta Cedergren (eds.): *Variation omnibus*. Edmonton, Alberta: Linguistic Research, pp. 115-124.

- Torres Cacoullos, Rena and Catherine E. Travis (2011): “Using structural variability to evaluate convergence via code-switching”, *International Journal of Bilingualism* 15, pp. 241–267.  
<http://dx.doi.org/10.1177/1367006910371025>
- Travis, Catherine (2007): “Genre effects on subject expression in Spanish: Priming in narrative and conversation”, *Language Variation and Change* 19, pp.101-135.  
<http://dx.doi.org/10.1017/S0954394507070081>
- and Rena Torres Cacoullos (2012): “What do subject pronouns do in discourse? Cognitive, mechanical, and constructional factors in variation”, *Cognitive Linguistics* 23, pp. 711-748.
- Van der Wel, Robrecht P.R.D., Robin Fleckenstein, Stephen Jax and David Rosenbaum (2007): “Hand path priming in manual obstacle avoidance: Evidence for abstract spatiotemporal forms in human motor control”, *Journal of Experimental Psychology: Human Perception* 33, pp. 1117-1126.
- Weiner, E. Judith and William Labov (1983): “Constraints on the agentless passive”, *Journal of Linguistics* 19, pp. 29-58.  
<http://dx.doi.org/10.1017/S0022226700007441>



**“Dude was figureando hard”:  
El cambio y la fusión de códigos  
en la obra de Junot Díaz**

| Domnita Dumitrescu  
| *California State University, Los Angeles*

**Abstract:** This article examines how the Dominican-American writer Junot Díaz uses Spanish in his Pulitzer prize-winning novel, *The Brief Wondrous Life of Oscar Wao*, (2007), and his most recent collection of short stories, *This Is How You Lose Her* (2012), both published in English. After reviewing the main claims made in different scholarly studies regarding code-switching (in general and in literature), the article elaborates on several comments from interviews with Díaz—who is bilingual, multidialectal, and controls a large spectrum of sociolinguistic registers—, in which he explains the reasons for which Spanish is an essential component of his work. Different types of code-switching and hybrid lexical creations, often humoristic, are discussed, and the conclusion is that, unlike his predecessors, Díaz does not limit himself to the use of a sustained alternation between Spanish and English, but rather, through congruent lexicalization, he produces a sort of code-fusion which is reflective of the heteroglossic practices of postcolonial bi- and multilingual societies, known as *translanguaging*.

**Keywords:** Code-mixing, bilingualism, congruent lexicalization, code-fusion, translanguaging.



## 1. INTRODUCCIÓN: EL CAMBIO DE CÓDIGO

El cambio de códigos (que se conoce en inglés como *code-switching*) es un fenómeno prevalente en el habla de los bilingües de cualquier parte del mundo<sup>1</sup>, y que oímos a diario en las conversaciones de muchísimos hispanos estadounidenses (o *hispanounidenses*, como se les suele llamar en la Academia Norteamericana de la Lengua Española), quienes alternan con desenfado el español y el inglés en sus interacciones verbales, con una amplia gama de propósitos comunicativos y efectos estilísticos, que mencionaré más tarde<sup>2</sup>.

El fenómeno del *code-switching* ha sido definido de diferentes maneras y ha sido estudiado en una amplia variedad de contextos, usando una gran diversidad de enfoques. Basta con solo recordar (en una mirada a vuelo de pájaro muy incompleta), que el fenómeno ha sido analizado en la conversación (Auer 1998; Lipski 1982; Zentella 1982, 1997), en las entrevistas sociolingüísticas (García 2009), en los encuentros de servicio (Callahan 2009), en las narrativas orales basadas en experiencias personales (Álvarez 1991; Koike 1987) o en el relato de los cuentos de hadas (Toribio 2000), y también en los textos literarios de todos los géneros (un tema sobre el cual voy a volver un poco más tarde). Además, ha sido abordado con una metodología estructuralista (Poplack 2000; Timm 1975), generativista (D’Introno 1996; Toribio y Rubin 1996; y también Jeff Mac Swanson en Bullock y Toribio, 2009), psicolingüística (Bullock y Toribio 2009), sociolingüística (Heller 1988; Lipski 1982, 2008), desde el ángulo de la teoría del *Matrix Language Frame* (Myer-Scotton 1993; Callahan 2002, 2004), desde el ángulo de la teoría de la marcidez (Mendieta, Lombard y Cintron 1995); desde la perspectiva cognitivo-pragmática del modelo lingüístico dual (Kecskes 2006); desde los parámetros de un modelo sincrético (Becker 1997); con una finalidad pedagógica en mente (Carvalho 2012; Sayer 2013); e incluso desde una óptica interdisciplinaria totalmente ecléctica (Rodríguez González y Parafita Couto 2012), por citar solo algunos ejemplos. Otros trabajos dignos de mencionar

---

<sup>1</sup> Por ejemplo, Franceschini considera que el cambio de código es un universal lingüístico en los ambientes multilingües: “There seems at this point to be more and more evidence that CS [*codeswitching*] is a language universal in the behaviour of multilingual speakers or —to employ a shorthand definition of CS— using several languages and language varieties in the course of a conversation is based on conversational-internal mechanisms observable in various social contexts all over the world” (1988: 51).

<sup>2</sup> Para información detallada con respecto a los fenómenos de contacto con el inglés en el español estadounidense, remito, en primer lugar, a los trabajos fundacionales de Carmen Silva-Corvalán (1983, 1989, 1994 y 2001); y también, entre muchos otros, a Dumitrescu (2011), Fuller (2013), Gimeno Menéndez y Gimeno Menéndez (2003), Klee y Lynch (2009), López-Morales (2009), Zentella (1982, 1997).



en esta línea son, por ejemplo Franceschini (1998), Pfaff y Chávez (1986) y Valdés-Fallís (1976 a). Como se puede notar, hay realmente una amplia variedad de enfoques tanto en lo que concierne al ámbito estudiado como también al marco teórico en que he llevado a cabo la investigación.

Por ello, me parece conveniente, para el propósito del presente trabajo, adoptar una de las más amplias definiciones descriptivas del *code-switching*, a saber la que propone Callahan en sus trabajos de 2004 y 2009:

Codeswitching is the use of words and structures from more than one language or linguistic variety by the same speaker within the same speech situation, conversation or utterance. Conversational codeswitching refers to the use of two languages by the same speaker within the same speech event. Codeswitching may occur at inter- and intrasentential levels, and may consist of single words or phrases. (2004:5)

Esta definición abarcadora permite, por una parte, englobar bajo una misma etiqueta fenómenos que algunos lingüistas han clasificado de diferentes maneras; por ejemplo Peter Muysken (2000) habla de *code-mixing* —mezcla de códigos— cuando se trata de alternar lenguas dentro del mismo enunciado (como en: “*Y luego after the war, ya no pudo regresar*”) y reserva el nombre de *code-switching* para la alternancia entre enunciados, o sea inter-oracional (como en: “*Ella no sabe hacerlo. I’ll do it*”)<sup>3</sup>. Por otra parte, hace posible, por razones descriptivas, dejar de lado la distinción que se ha tratado de establecer a veces entre un cambio de código que consiste en una sola palabra y un préstamo léxico no adaptado. Finalmente, ofrece una buena transición hacia lo que he llamado “fusión de códigos” en ciertas obras literarias, como las del autor que forma el objeto de este estudio (cf. Dumitrescu 2013c), y que expondré a continuación, en la última sección del trabajo, en conexión con el concepto del “translenguar”.

## **2. EL CAMBIO DE CÓDIGO ENTRE EL ESPAÑOL Y EL INGLÉS Y EL LLAMADO SPANGLISH**

El cambio de códigos conversacional entre hispanounidenses (mezclado o no con préstamos adaptados y con calcos léxico-semánticos

<sup>3</sup> Ejemplos oídos por la autora en Los Ángeles, California.

del inglés) se conoce popularmente como *Spanglish* (o “espanglish” en la forma hispanizada en que aparecerá en la próxima edición del *DRAE*). No es mi intención volver a exponer aquí la polémica suscitada por este concepto, ya que, siguiendo a Otheguy y Stern 2010 y Lipski 2008, la he discutido en otros artículos míos (e.g., Dumitrescu 2010, 2012a, 2013 a,b). Basta con mencionar que el término se presta a confusiones, por lo que no lo voy a usar en el presente trabajo, aun cuando haya críticos literarios que sí lo han empleado refiriéndose a los autores hispanounidenses, como el que voy a abordar en este estudio. Por ejemplo en *Weird English* (2004), Evelyn Ch’ien dice que Junot Díaz, en sus obras, hace un uso del *Spanglish* “overtly courageous”, o en el *New York Times* se afirma que el autor usa “a sort of streetwise brand of *Spanglish* that even the most monolingual reader can easily inhale” (Kakutani 2007). Asimismo, generalizando, Silvia Betti afirma que “es en la literatura [...] donde el *Spanglish* adquiere su sello de legitimidad dando a ese lenguaje autoridad y verosimilitud, gracias a escritores chicanos, puertorriqueños, cubanoamericanos, dominicanos, entre otros” (2008: 115).

Pero lo que sí quiero destacar es lo siguiente: una conclusión unánime a la que han llegado los investigadores del intercambio de códigos (o como queramos llamarlo) es que los hablantes no mezclan las lenguas al azar, y no lo hacen por no dominarlas en suficiente medida, sino que, al contrario, siguen unas reglas sistemáticas al hacerlo (ver Poplack 2000 al respecto, y también, entre otros, Pfaff & Chávez 1986 y Timm 1975), y sus motivaciones están lejos de originarse en la ignorancia o pereza mental (como es común pensar). Acerca de esta confusión, Bullock y Toribio (2009) declaran lo siguiente (en la introducción a su volumen editado):

Broadly defined, C[ode] S[witching] is the ability on the part of bilinguals to alternate effortlessly between their two languages. This capacity is truly remarkable and invites scientific and scholarly analysis from professionals, but, at the same time, generates a great deal of pointed discussion that reflect popular misperceptions of the nature of CS in particular and bilinguals more generally. While CS is viewed as an index of bilingual proficiency among linguists, it is more commonly perceived by the general public as indicative of language degeneration. (2009: 1)

Y más adelante, las mismas autoras reiteran esta idea cuando afirman:

[A] significant body of research has amply demonstrated that CS does not represent a breakdown in communication, but reflects the skillful manipulation of two language systems for various communicative functions. This is articulated through a different –and to our mind, more apt–metaphor offered by Valdés: “[I]t is helpful to imagine that when bilinguals code-switch, they are in fact using a twelve-string guitar, rather than limiting themselves to two six-string instruments” (1988:126). (Bullock y Toribio, 2009: 4-5)

Efectivamente, hay mucha evidencia que apunta hacia el hecho de que son precisamente los hablantes bilingües con más dominio de las dos lenguas quienes las alternan con más frecuencia, aparentemente con el deseo, consciente o inconsciente, de afirmar su pertenencia simultánea a dos culturas y de dar voz a su identidad dual. Por ejemplo, Silva-Corvalán afirma:

Como lo han demostrado numerosos estudios, la alternancia de lenguas es un estilo comunicativo característico de grupos con niveles más o menos equilibrados de competencia bilingüe. Se da de manera limitada en otros grupos, para quienes la alternancia tiene más bien un valor emblemático o sirve para compensar el olvido o la falta de material lingüístico. Por otra parte, la función retórica de la alternancia es manejada con maestría por los bilingües equilibrados. (2001: 317)

Desde luego, en esta cita se alude al importante problema de la distinción entre los bilingües equilibrados y aquellos hispanounidenses que están sufriendo el proceso llamado en inglés *language attrition*, que viene a ser, en esencia, una pérdida gradual de su competencia comunicativa en la lengua de herencia, lo cual los hace acudir al inglés como estrategia compensatoria, para poder hacerse entender<sup>4</sup>. A estos hablantes, Lipski

<sup>4</sup> Es interesante que Silva-Corvalán haya observado esta distinción hace tiempo, en un estudio de 1983 en el que propone acertadamente distinguir entre *code-switching*, en el sentido habitual que se le atribuye a la palabra, y lo que ella llama *code-shifting*, o sea precisamente lo que hacen los hablantes que están perdiendo su lengua, al acudir a la lengua que mejor dominan, a fin de hacerse entender. He aquí las palabras de Silva-Corvalán al respecto: “Code-switching and code-shifting may be two different phenomena. Code-switching appears to require a large degree of linguistic competence in the two languages and is largely motivated by social and discourse/pragmatic factors. Code-shifting, on the other hand, fulfills basically a linguistic function and is motivated by a specific sociolinguistic situation: the need to communicate in a language in which the speaker has a limited degree of competence” (1983: 85).

(2008) los llama “bilingües de transición” o “vestigiales” y lamenta el hecho de que muchas veces se les haya considerado representativos del español estadounidense en su conjunto<sup>5</sup>.

No me estoy refiriendo, pues, a ese grupo, sino a aquellos hispanounidenses que hablan español con fluidez, pero que, al vivir insertos en la cultura norteamericana, resuelven este desfase recurriendo a la convergencia cultural y produciendo lo que Ricardo Otheguy llama “un discurso culturalmente norteamericanizado”, pero detrás del cual no se halla “una lengua estructuralmente anglicada” (Otheguy 2013: 132). Para estos hablantes, la alternancia de códigos parece ser la mejor forma de validar lingüísticamente su identidad dual, y no pocas veces se transforma en una verdadera necesidad vital de la que depende su autoestima y su legitimidad socio-cultural. Valga al respecto el testimonio de la conocida escritora chicana Gloria Anzaldúa:

Until I can accept as legitimate Chicano Texas Spanish, Tex-Mex and all the other languages I speak, I cannot accept the legitimacy of myself. Until I am free to write bilingually and to switch codes without having always to translate, while I still have to speak English or Spanish when I would rather speak Spanglish, and as long as I have to accommodate the English speakers rather than having them accommodate me, my tongue will be illegitimate. (1999: 59)

### 3. EL CAMBIO DE CÓDIGO EN LA LITERATURA Y SU RAZÓN DE SER

Como mencioné antes, hay una profusión de estudios sobre el cambio de código no solo en la conversación, sino también en la literatura hispanounidense. La mayoría, como es de esperar, se refieren a la literatura chicana, que fue la primera en adoptar este estilo comunicativo como recurso estilístico propio (cf. Silva-Rodríguez y Hernández G. 2012), pero también hay un número de estudios, sobre todo recientes, acerca de autores

---

<sup>5</sup> Según Lipski, “[T]ransitional bilinguals—most of whom are regarded, and regard themselves, as true bilinguals—in business, politics, journalism, law enforcement, and the arts are frequently used as examples of U.S. Latino Spanish speakers, and much of the criticism directed at Spanglish as an impoverished language spoken in the United States stems from confusing the symptoms of transgenerational language attrition with stable bilingualism” (2008: 64).

hispanounidenses de otras etnicidades (en particular, de origen caribeño). Para citar otra vez a vuelo de pájaro algunos, recordemos, para la poesía, las contribuciones de Flores (1987), Keller & Randall (1993), Mendieta-Lombardo y Cintron (1995), Timm (2000), Valdés Fallís (1976b, 1977); para el teatro, Jonsson (2010); y, para la prosa, Bürki (2003, 2008), Callahan (2003, 2004), Derrick (2014), Keller (1984), Lipski (1982), Montes-Alcalá (2012), Rudin (1996), Torres (2007). También ha sido abordado el problema de la traducción de los textos con cambio de código (e.g., García Vizcaíno 2008 y Díaz Pérez 2012) y se han escrito no pocas tesis doctorales sobre este fenómeno en los textos literarios actuales (ver, por ejemplo, Burrows 2010; Hansen Esplín 2012; Montes-Alcalá 2000). El consenso entre los especialistas es que el cambio de código en literatura, en vez de considerarse un déficit de los textos que lo incorporan, representa una innovación digna de encomio. Como se expresa Aparicio:

While some prescriptive linguists, editors, and authorities in education would judge the interference of Spanish and English as a deficit, a postmodern and transcreative approach would validate it as a positively creative innovation in literature. (1994: 797)

Varios investigadores, a lo largo del tiempo, han intentado describir el papel artístico del cambio de códigos en la literatura. Uno de los pioneros de este tipo de estudios es, por ejemplo, Garry Keller, quien distingue entre un cambio de códigos temático, un cambio de códigos para la caracterización y un cambio de códigos con funciones estilísticas (cf. Keller 1984 y Keller y Randall 1993). Otros, más tarde (como Jonsson 2010, por ejemplo), han privilegiado diferentes enfoques analíticos, que el espacio no me permite abordar. Por ello, prefiero acudir a una excelente síntesis reciente acerca de las más significativas contribuciones al análisis del cambio de código en literatura, a saber el artículo de Cecilia Montes-Alcalá, “Code-switching in US-Latino novels” (2012), donde se estudian las funciones socio-pragmáticas del cambio de código en un corpus bilingüe compuesto por tres novelas: una chicana (*Mi querido Rafa*, por Rolando Hinojosa), una cubano-americana (*La vida es un Special*, por Roberto Fernández) y una nuyoricana (*Yo-Yo Boing!*, por Giannina Braschi).

Estas funciones, en la formulación de Montes-Alcalá, son las siguientes: 1. *Lexical need*, o sea un cambio al nivel de la palabra, debido al biculturalismo del hablante, que no encuentra un equivalente exacto en

la otra lengua para elementos léxicos cargados de connotaciones culturales y/o emocionales; 2. *Clarification and elaboration*, o sea un cambio de idioma destinado a elaborar sobre algo dicho previamente, o a clarificar, a través de repeticiones o comentarios parentéticos, un concepto de la otra lengua; 3. *Stylistic code-switching*, usado, por ejemplo, para reflejar el habla coloquial de una forma realista y creíble, o en oraciones que se pueden atribuir al “fluir de conciencia”; 4. *Idioms, discourse markers and linguistic routines*, o sea inserciones de expresiones fijas y de marcadores discursivos de la otra lengua, a veces incluso de forma inconsciente; 5. *Emphasis*, o sea un cambio de código destinado a llamar la atención, a enfatizar, una palabra o una idea, que corresponde a lo que otros han llamado *foreground* —‘poner en primer plano’—; 6. *Quotations*, o sea el uso de la lengua original para citar las palabras de otra persona, como discurso directo o parafraseándolas; 7. *Triggered switches*, o sea cambios de idioma provocados por un cambio que le precede o le sigue (como por ejemplo lo que se ha dado en llamar *anticipatory triggering*) (Montes-Alcalá 2012: 75). (Hay que tener en cuenta, sin embargo, que un cambio puede realizar más de una función a la vez, y también que hay cambios cuya función puede no ser tan transparente).

El propósito del estudio bajo discusión es averiguar si estas siete funciones socio-pragmáticas identificadas hasta ahora en la alternancia de códigos que ocurre en el discurso bilingüe natural se da también en el discurso bilingüe literario (en otras palabras, si este discurso es mimético, es un reflejo del habla de la comunidad bilingüe); y las conclusiones del estudio —realizado con métodos estadísticos rigurosos— ofrecen una respuesta afirmativa. En otras palabras, según Montes-Alcalá, el cambio de código en la literatura “may be considered authentic and not just purely rhetorical”, y juega, sin duda “a crucial role in contemporary US-Latino literature”. La autora del estudio concluye afirmando que “[t]hese writers’ language choice reveal how bilingual individual living in the hyphen —between two cultures— can and sometimes need to write in both languages in order to fully express themselves” (2012: 86).

Efectivamente, lo que al parecer distingue a los autores hispanounidenses de otros escritores que usan más de una lengua en sus creaciones exclusivamente como procedimiento estilístico (la *Renga* de Octavio Paz sería un ejemplo de ello- cf. Dumitrescu 2014), es el hecho de que, al acudir al cambio de código (o *Spanglish*, como dicen ellos), transfieren a la literatura la realidad lingüística de la expatriación o de la marginación

social tal como la viven millones de personas en Estados Unidos (u otros países donde su lengua nativa no es la mayoritaria, como por ejemplo los francófonos del Canadá— cf. Ladouceur 2006). Como señala Torres también, hay consenso entre los expertos acerca del hecho de que el cambio de código es “una opción artística *con ramificaciones políticas*” (2007: 76, traducción y énfasis mío). Y agrega Torres:

[C]ode-switching in literature is not only metaphorical, but represents a reality where segments of the population are living between cultures and languages; literary language actualizes the discourse of the border and bilingual/bicultural communities. (2007: 76)

Es más: después de citar a Mónica Heller, para quien “[c]ode switching is a microlevel symptom of a macrolevel change” (1988: 12), Callahan explica (y comparto su opinión) que

[i]ts use by authors —whether or not this is their conscious intention— re-defines social reality; and influences what discourse resources are available for what domains. This includes a rejection of monolingual English as well as of monolingual Spanish. (2004:145)

En otras palabras, estos autores “challenge the balance of power in the language marketplace in the United States. ... These writers’ use of Spanish/English codeswitching in their work is a *de facto* claim to the right to participate in the marketplace on a bilingual and not just monolingual basis” (2004: 4).

#### 4. JUNOT DÍAZ Y EL PAPEL DEL ESPAÑOL EN SU PROSA

Voy a ilustrar las consideraciones anteriores comentando la obra del narrador hispanounidense que más interés ha despertado recientemente, especialmente después de haber sido galardonado hace unos años con el Premio Pulitzer de Literatura para 2008 y haber encabezado durante semanas la lista de best-sellers del *New York Times*: el dominicano-estadounidense Junot Díaz (n. 1968), autor, hasta la fecha, de tres libros: *Drown*, 1997 (una colección de cuentos traducida al español como *Negocios*); *The Brief Wondrous Life of Oscar Wao*, 2007 (novela traducida al español el año



siguiente como *La breve y maravillosa vida de Óscar Wao*); y *This is How You Lose Her*, 2012 (colección de novelas cortas traducidas al español en mayo de 2013, con el título *Así es como la pierdes*)<sup>6</sup>.

De la prosa de este “chico terrible de la literatura estadounidense” (como se le llamó en una entrevista con Eduardo Lago en *El País Semanal*, 5 de mayo de 2013)—que vive en Nueva York, es catedrático de escritura creativa en el famoso Massachusetts Institute of Technology y ha recibido un número impresionante de reconocimientos—Ofelia García ha dicho, y creo que con mucha razón, que “parece indicar un cambio epistemológico en la presencia del español en los Estados Unidos” (2013: 368).

Para empezar, Díaz, a diferencia de otros autores hispanounidenses, intercala en inglés—la lengua en que escribe sus obras— vocablos y frases en español sin comillas, cursivas u otro signo ortográfico que los identifique como elementos foráneos<sup>7</sup>. Y lo hace con un propósito bien claro, que explica en una entrevista de 2000 con Silvio Torres-Saillant y Diógenes Torres Céspedes:

For me allowing the Spanish to exist in my text without the benefit of italics or quotations marks a very important political move. Spanish is not a minority language. Not in this hemisphere, not in the United States, not in the world inside my head. So why treat it like one? Why ‘other’ it? Why de-normalize it? By keeping Spanish as normative in a predominantly English text, I wanted to remind readers of the fluidity of languages, of the mutability of languages. And to mark how steadily English is transforming Spanish and Spanish is transforming English. (citado en Ch’ien, 2004: 204)

También vale la pena citar este otro pasaje de la misma entrevista, donde Díaz explica el porqué de la incorporación de tantas voces española en sus textos ingleses:

<sup>6</sup> Hay que observar que los estudios sobre su obra no han tardado en aparecer incluso cuando era autor de una sola obra, *Drown* (cf. Frydman 2009, Kevane 2003), lo que demuestra que desde el principio se le percibió como una voz insólita y potente dentro del panorama literario norteamericano; tampoco han faltado tesis doctorales, como la de Baez (2008), que se enfoca en el tema del exilio entre los autores dominico-americanos.

<sup>7</sup> Aparentemente, al principio Díaz mismo tuvo que usar cursivas en los pasajes en español para satisfacer el requisito del editor de su cuento “The Sun, the Moon, and the Stars”, pero después de esto estipuló en todos sus contratos que los editores debían aceptar sus obras sin cursivas, o si no, no las podían publicar (cf. Ch’ien, 2004: 207). Otra escritora que usa palabras españolas sin cursivas en su texto es Susana Chávez-Silverman (cf. Derrick 2014).



Look for instance at dictionaries and their way of adopting foreign words. When does a loan word become an English word? Is “hacienda” a word in Spanish or English? You know what I’m saying? The point is, I am pushing the dates on a lot of these words. I decided I don’t need a hundred years for the *Oxford English Dictionary* to tell me that it’s okay to adopt this or that word as part of our normal vocabulary. I feel that’s what we always should do. We should be pushing the dates on words. (citado en Ch’ien, 2004: 204)

Según él mismo confiesa, el español y el inglés han estado en conflicto en su conciencia, y la lucha por sobrevivir del primer idioma ante el segundo, aparentemente más poderoso (al menos a los ojos del futuro autor cuando todavía era niño y adolescente) se refleja también en este estilo suyo específico de idiomas mezclados. Por una parte, Díaz —quien llegó a los Estados Unidos a los seis años de edad— declara lo siguiente acerca de su dominio del español (en la ya citada entrevista que le hicieron en *El País* en 2013):

Quando llegué a este país perdí el español y de adulto dediqué un gran esfuerzo a intentar recuperarlo. La sensación de pérdida es muy aguda y dolorosa. El inglés siempre ha sido una sombra gigantesca que se cernía sobre el español, pero también es cierto lo contrario. La sombra del español se cierne sobre el inglés. Para mí, hablar bien español siempre ha sido una obsesión, pero lo cierto es que mi español es mediocre, cosa que genera en mí un enorme sentimiento de culpa. Por otra parte, cuando hablo español soy otro, mi personalidad es distinta, más ligera. (Lago 2013)

Por otra parte, en la entrevista con Silvio Torres-Saillant y Diógenes Torres Céspedes (citada tanto en Ch’ien, 2004 como en Frydman, 2009), declara lo siguiente:

When I learned English in the States, this was a violent enterprise. And by forcing Spanish back into English, forcing it to deal with the language it tried to exterminate in me, I’ve tried to represent a mirror image of that violence on the page. Call it my revenge on English. (citado en Frydman, 2009:134)

Para demostrar cómo usa este autor el español en su obra (o como “toma su venganza” en contra del inglés, para aludir a la cita anterior), voy a comentar a continuación algunos ejemplos sacados de su novela premiada *The Brief Wondrous Life of Oscar Wao*, a la que el periodista que le entrevistó para *El País* considera “una obra que a pesar de estar escrita en inglés lleva en sí las señas de identidad de una cultura en español que ha cambiado de manera irreversible el rostro de Estados Unidos”, y también algunos ejemplos de su último libro, *This Is How You Lose Her*<sup>8</sup>.

Lo que quiero plantear, en última instancia, es que Junot Díaz, además de usar el cambio de código como han hecho, antes de él, otros escritores hispanounidenses (por ejemplo, Sandra Cisneros, Oscar Hijuelos y otros, que fueron sus maestros, según él mismo declaró una vez), va más allá de sus antecesores y le propone al lector una fusión ingeniosa de ambos idiomas, que prácticamente no tiene precedente<sup>9</sup> y que representa lo más innovador y original de su creación literaria. Adoptando la terminología de García (2013) y García y Wei (2013), considero esta fusión de códigos en la obra del escritor dominicano-americano como una forma literaria de translenguar, un tema que voy a desarrollar en la última parte de este trabajo.

<sup>8</sup> En las citas que siguen, estas obras aparecen abreviadas como, respectivamente, *Oscar* y *Lose*. Ambas obras están traducidas al español, como se ha dicho, pero, a mi modo de ver, muchos matices se pierden en la traducción, a pesar de que Díaz piensa que éste no sea el caso. Por ejemplo, a la periodista de *La Nación* que le preguntó si “no se desluce un poco su prosa al perderse en parte la cruza de inglés y castellano, la potencia expresiva que produce la combinación”, Díaz contestó lo siguiente: “No, porque el inglés y el castellano viven juntos. Son como primos amargados, cada uno conoce al otro hace muchísimos años. Como dice un amigo mío: ‘Siempre están metidos en la misma cama’. En la versión en español hay muchísimas cosas en inglés y se mantiene el mismo equilibrio entre los dos idiomas. Ese pleito, ese amor que tienen. Yo creo que si tuviera que traducirlo a otro idioma —al ruso, por ejemplo— sería más difícil. El inglés y el español son mellizos (aunque no quieran reconocerlo) y fueron separados al nacer. Tal vez alguien de otro lado no lo vea de esta manera, pero aquí, en esta parte del mundo, resulta bastante claro” (Basavillbaso 2013).

<sup>9</sup> Las escritoras que más se acercan, desde este punto de vista, a Junot Díaz (sin haber logrado, sin embargo, su notoriedad, y sin haber publicado tanto) son, aparentemente, Giannina Braschi y Susana Chávez-Silverman, a quienes Torres (2007) incluye en la categoría de los escritores hispanounidenses que usan un *sustained code-switching* a causa del cual solo pueden ser leídos por un público bilingüe. Torres considera que este tipo de textos que experimentan con la mezcla de idiomas hasta el punto de ser prácticamente ininteligibles para un lector monolingüe exhiben un “bilingüismo radical”, a diferencia de otros que usan estrategias diferentes (y desde luego menos “radicales”) para incluir el español en textos ingleses (como por ejemplo el uso de vocablos españoles fáciles de entender en el contexto, o la traducción con valor explicativo). Pero, como se verá más adelante, hay una diferencia importante entre la escritura de estas autoras y el tipo de prosa que produce Díaz.

## 5. CAMBIO DE CÓDIGOS, CAMBIO DE REGISTROS Y CAMBIO DE DIALECTOS EN LA OBRA DE JUNOT DÍAZ

Casi todas las funciones socio-pragmáticas identificadas por Montes-Alcalá (2012) para el cambio de códigos tanto en la interacción oral espontánea como en los textos literarios (y comentadas brevemente en la sección 3, arriba), están presentes, en diferentes grados, en las obras de Junot Díaz. Pero las cuatro que, a mi modo de ver prevalecen son: la reproducción de citas directas, el uso estilístico y/o enfático, la inserción de expresiones idiomáticas y las necesidades léxicas (en el caso de conceptos y términos que conllevan una carga cultural idiosincrásica). Los ejemplos a continuación van a ilustrar todas estas funciones.

En primer lugar —y lo mismo que en otros autores hispanounidenses que acuden al *code-switching*—, el español —solo o alternando con el inglés en el mismo turno conversacional— se emplea casi sistemáticamente para reproducir el discurso directo de los personajes hispanos. Por ejemplo:

- (1) Que [*sic!*] muchacha tan fea, she said in disgust. (*Oscar* 54)
- (2) Abuela snorted. Guapa soy yo. Your mother was a diosa. But so cabeza dura. (*Oscar* 75)
- (3) Pedro snorts: Ese ladrón no va' pa' ningún la'o. (*Oscar* 112)
- (4) Sounds like you're going to be bien cómoda, Marisol says. (*Lose* 65)
- (5) Look at that little macho, his mother's friend said. Que [*sic!*] hombre (*Oscar* 14)
- (6) You remember ...how skinny she was, no culo, no titties, como un palito, but your brother didn't care. (*Lose* 148)
- (7) And since her mother was una maldita borracha (to quote Oscar's mom)... (*Oscar* 13)

Este discurso directo, muchísimas veces, incluye palabras malsonantes, interjecciones vulgares e insultos de diversos tipos, ya que las más de las veces los diálogos ocurren en situaciones “explosivas”, de conflicto o de confrontación verbal entre los interlocutores, y el autor acude en estos casos a un cambio al registro informal. Como señala Callahan, “it is not unusual for codeswitching to co-occur with taboo words, which are an unmistakable marker of a register associated with an informal style. ...[in which] expletives and other interjections are a prime locus for switching” (2003:15). He aquí unos cuantos de los muchos ejemplos de este tipo que se pueden sacar de la prosa de este escritor:

- (8) Pa'fuera... What in carajo is the matter with you? (*Oscar* 22)
- (9) Hija de la gran puta, would you stop jodiéndome! (*Oscar* 130)
- (10) Your mother flares: Who in carajo do you think you are talking to? Say hello, coño, to la profesora. (*Lose* 155)
- (11) You, stupid, worthless no-good hijo-de-la-gran-puta are going home. (*Oscar* 302)
- (12) I said, Rafa, un chín de respeto .What the fuck do you know? Nothing.  
Nothing at all. Entonces cállate la fucking boca. (*Lose* 94)
- (13) Figurín de mierda, she called me. You think you are somebody, but you ain't nada. (*Oscar* 60)

Sirva de ejemplo adicional este diálogo guasón entre Rafa y su madre, acerca de una vecina con la cual el muchacho se había acostado:

- (14) What do you want, Ma? Se metió por mis ojos.  
Por mis ojos my ass, she had said. Tú te metiste por su culo.  
That's true, your brother admitted cheerily. Y yo por su boca.  
(*Lose* 156)

Se puede conjeturar que la función pragmática principal de estos términos “tabú”, dentro de las citas directas, es la de recalcar, de dar énfasis, a los actos de habla expresivos realizados en los enunciados que los contienen, y, al mismo tiempo, crear efectos humorísticos (funciones que, por lo demás, han sido comentadas por García Vizcaíno 2008 en su análisis del intercambio de códigos en Sandra Cisneros). Sin embargo, hay que reconocer que no todo es vulgar y violento en el habla de los personajes. A veces hay pasajes llenos de ternura, como este, por ejemplo, donde la narradora recuerda las caricias verbales que le prodigaba su amante:

- (15) He made her feel guapa and wanted and safe....On their nights together, he would pass his hand over her naked body...murmuring, Guapa, guapa, over and over again. (He didn't care about the burn scars on her back: It looks like a painting of a ciclón and that's what you are, mi negrita, una tormenta en la madrugada). (*Oscar* 127)

Pero lo que es más interesante observar es que, la mayoría de las veces, Díaz, al citar los discursos de sus personajes hispanos, no recurre tanto a los “clásicos” cambios de códigos inter- o intra-oracionales aislados, como

en los ejemplos arriba citados, sino que suele intercalar de forma repetida el español en el inglés en un solo y mismo enunciado (o serie de enunciados pronunciados por el mismo interlocutor), creando un constante “pendular” entre una lengua y la otra, lo que potencia al máximo la expresividad (y la autenticidad) de los respectivos pasajes. He aquí, a título ilustrativo, los cínicos, aunque bien intencionados, consejos que el experto en “asuntos sexuales” tío Rudolfo le da a su sobrino, que es todavía virgen a una edad a la que otros han dejado de serlo desde hace tiempo:

- (16) Listen, palomo, you have to grab a muchacha y méteselo.  
That will take care of everything. Start with a fea. Coje [*sic!*]  
that fea y méteselo ... Hijo, you are the most buenmoso[*sic!*]  
man I know. (*Oscar* 24)

Observen también esta reacción airada de Beli, ante las amenazas de la hermana de Trujillo, que le informa que le convendría que jamás se volviera a topar con su “black cara de culo” si no quiere ir a la cárcel:

- (17) Beli ...still had the ovaries to spit, Cómeme el culo, you ugly  
disgusting vieja! (*Oscar* 141)

Voy a volver más adelante sobre este tema de los enunciados que contienen cambios idiomáticos repetidos.

Por otra parte, y también como han hecho muchos otros autores hispanounidenses antes de él, Junot Díaz usa el español para introducir dichos populares, modismos, o para insertar una serie de términos dialectales (en este caso, del habla dominicana popular) o del barrio, que confieren autenticidad y sabor al texto. He aquí algunos ejemplos de dichos o refranes:

- (18) As the viejos say, clavo saca clavo. (*Oscar* 124)  
(19) Like they say: plátano maduro no se vuelve verde. (*Oscar* 208)  
(20) Like they say: los que menos corren, vuelan. (*Oscar* 196)

O un ejemplo de modismo (“salírsele a uno el lobo”):

- (21) How about I buy you a drink? he said, and when she turned  
away como una ruda, he grabbed her arm, hard, and said,  
Where are you going, morena? And that was all it took: a  
Beli le salió el lobo... .When baller twisted her arm, she went  
from zero to violence in under .2 seconds. Shrieked: *No.*  
*Me.Toques.* (*Oscar* 115)

Al lado de los refranes y modismos, algunos autores incluyen en esta categoría de cambios de códigos el uso de marcadores discursivos y de términos relacionados directamente con la identidad de los hablantes. En este sentido, creo que el hecho de que Junot Díaz usa siempre los términos de parentesco en español es una prueba de que este idioma es el idioma que expresa mejor sus raíces y las de sus personajes (muchos de los cuales son en realidad alter-egos del autor)<sup>10</sup>. De hecho, Junot Díaz parece identificar el español con la lengua de la familia, de la intimidad, de la confianza, y sugiere que ser tratado en esta lengua es un privilegio que uno debe ganar (y puede perder, si su conducta es impropia). Esto se puede ver en el siguiente pasaje, que contiene el único caso en que se usa *father* en vez de “padre”, seguramente para marcar el distanciamiento que ocurre entre los dos hombres, después de la ruptura del narrador con su novia.

(22) Her father, who used to treat me like his hijo, calls me an asshole on the phone, sounds like he’s strangling himself with the cord. *You no deserve I speak to you in Spanish, he says* (énfasis mío). (*Lose 4*)

En cuanto a los términos con connotaciones socio-culturales específicas, o para los cuales el autor considera que no hay equivalente expresivo exacto en la otra lengua, he aquí algunos ejemplos que contienen dominicanismos (a menudo del registro vulgar), como *popola* y *chocha* (órgano genital femenino), *fukú* (maldición), *ripio* (pene) y *jojote* (albino), entre otros<sup>11</sup>:

<sup>10</sup> Es interesante en este sentido pensar en el uso de los términos de parentesco que hace Richard Rodríguez en su famosa autobiografía de 1982, *Hunger for Memory*, donde explica cómo, después de que el inglés se convirtió en su primera lengua, ya no supo qué palabras usar para dirigirse a sus padres, ya que las antiguas palabras españolas que había usado antes —*papá* y *mamá*— ya no las podía usar; habrían sido unos recuerdos demasiado dolorosos de lo mucho que había cambiado en su vida. Por otra parte, las palabras con que oía a los chicos del vecindario llamar a sus padres, le parecían inapropiadas, no eran realmente términos con que dirigirse a sus padres. Como resultado, nunca los usó en casa. Siempre que hablaba con sus padres, trataba de llamarles la atención solo con la mirada (1982:23).

<sup>11</sup> En realidad, la lista de dominicanismos empleados en la obra de Junot Díaz es muy extensa, y tengo entendido que sus primeros escritos salían acompañados de un glosario (pero ahora ya no, ya que al escritor realmente no le interesa —como lo ha declarado repetidas veces— alcanzar la inteligibilidad absoluta entre sus lectores). Algunos ejemplos de términos y expresiones dialectales —relacionados con la realidad socio-cultural dominicana o dominicano-estadounidense— son: *pana* (como forma de tratamiento), gentilicios como *azuano*, *banilejo*, términos que describen características raciales (*cocolo*, *jíbara*, *jabao*, *pañol*), vestuario (*chacabana*), insultos típicos (*pocohombre*, *mamahuevo*), y muchos otros vocablos del habla diaria como *tíguere*, *tutumpote*, *jipeta*, *pariguayo*, *comparón*, *chupabarrío*, *quemagoma*, *ciguapa*, *bochinchera*, *manganzón*, *figurear*, *un chin*, etc., cuyo significado a veces ni siquiera un diccionario de americanismos como el de la Real Academia Española esclarece por completo. Agradezco a Porfirio Rodríguez su ayuda con los dominicanismos del texto.

- (23) Then, like now, Santo Domingo was to popóla [*sic!*] what Switzerland was to chocolate. (*Oscar* 121)
- (24) What Kennedy's intelligence experts failed to tell him is what every single Dominican, from the richest jabao in Mao to the poorest güey in El Buey, from the oldest anciano sanmacorisano to the littlest carajito in San Francisco knew: that whoever killed Trujillo, their family would suffer a fukú so dreadful it would make the one that attached itself to the Admiral jojote in comparison. (*Oscar* 3)
- (25) It's a well documented fact that in Trujillo's DR if you were of a certain class and you put your cute daughter anywhere near El Jefe, within the week she'd be mamando his ripio like and old pro and there would be nothing you could do about it! (*Oscar* 217)
- (26) I am thinking about Magda, how I'll probably never taste her chocha again. (*Lose* 23)

En una entrevista con Evelyn Ch'ien del 26 de julio de 2002, Junot Díaz, refiriéndose a su primer libro, *Drown*, y al hecho de que lo que más ha llamado la atención en él ha sido el uso del español, había declarado:

I just felt there were a number of registers in there—there was Dominican Spanish, general Spanish which is like a Spanish drawn from growing up with a bunch of Latinos, so it's not specific to Dominican-ness, then there is the various Englishes. I just found it really weird that the other language everyone kept focusing on was the homogenized Spanish. (citado en Ch'ien, 2004: 202)

Como se puede ver, esta mezcla no solo de idiomas y de registros formales, informales y vulgares (de ambas lenguas), sino también de variedades diatópicas (i.e., regionales) y diastráticas (i.e., sociales) del español continúa siendo una marca distintiva del estilo literario de Junot Díaz, que se puede caracterizar no solo... sino también como bilingüe, sino también como bidialectal (o multidialectal, si contamos también con las variedades de inglés que emplea en sus escritos, según menciona en la cita anterior).

Además, en relación con la "conciencia sociolingüística" del autor, por así decir, vale la pena señalar que Junot Díaz, en su afán de crear efectos de máxima autenticidad, incluye en sus escritos una serie de referencias



metalingüísticas a la variedad regional y al sociolecto usado por sus personajes. Por ejemplo, Constantina es caracterizada como “a new girl... whose cuerpo was all pipa and no culo, una ‘mujer alegre’ (*in the parlance of the period- énfasis mío*)”. (*Oscar* 113).

Otro ejemplo, en que se describe el habla popular de una dominicana estadounidense, que no ha perdido su acento regional, a pesar de llevar años en Quisqueya (el barrio neoyorkino Washington Heights, habitado principalmente por los inmigrantes dominicanos):

- (27) Those of you who have stood at the corner of 142<sup>nd</sup> and Broadway can guess what it was she spoke: the blunt, irreverent cant of the pueblo that gives all dominicanos cultos nightmares on their 400-thread-count sheets and that La Inca had assumed had perished along with Beli’s first life in Outer Azua, but here it was so alive, it was like it had never left: Oye, parigüayo [*sic!*], y qué pasó con esa esposa tuya? Gordo, no me digas que tú todavía tienes hambre? (*Oscar* 108)

Y finalmente este ejemplo, sobre el español dominicano rústico:

- (28) A total campesina from the way she held herself down to the way she talked, which was so demotic I could’t understand half of what she said– she used words like deguabinao and estribao on the regular. (*Lose* 101)

Es más: Díaz, como un verdadero dialectólogo aficionado, también recurre a lo que se ha llamado en la literatura de especialidad el “dialecto al ojo” o *eye dialect* (cf. Nuessel 2000 o Dumitrescu 2012b, entre otros), haciendo transcripciones literales de la pronunciación de sus personajes, como cuando usa *Nueva Yol* (*Oscar* 72), para Nueva York – con la neutralización de líquidas en final de sílaba o palabra, típicamente caribeña–; o cuando reproduce la elisión de la *d* intervocálica (como en *jabao*) y de la *s* en posición implosiva, también típicas, ambas, del Caribe y de otras hablas populares hispánicas. He aquí dos ejemplos para la *-d-* intervocálica elidida:

- (29) Every neighborhood has its tetúa, but Beli have put them all to shame, she was La Tetúa Suprema. (*Oscar* 92)
- (30) He was still the passionate enamorao who fell in love easily and deeply. (*Oscar* 23)



Y dos ejemplos para la -s implosiva elidida en la partícula evidencial *dizque*, muy frecuente en los textos dominicanos (cf. Dumitrescu 2012 c):

- (31) Despite “discovering” the New World the Admiral died miserable and syphilitic, hearing *dique* divine voices. (*Oscar* 1)  
 (32) I ran off, *dique*, because of a boy. (*Oscar* 61)

Asimismo, vale la pena mencionar que Junot Díaz capta en su prosa también otros fenómenos dialectales, típicos del habla caribeña, como por ejemplo, las preguntas parciales sin inversión, o la aféresis del verbo *estar*. He aquí unos ejemplos de aféresis de *estar*: “Doña, es verdad que tu hijo taba rapando una vieja?” (*Lose* 170); “When Oscar whimpered, Moms de León nearly exploded. Tú ta llorando por una muchacha?” (*Oscar* 14); “A neighbor lady with a congested daughter. Ta muy mal, she says. Of course everybody has a sister or a prima they want you to meet. Que tan mas [*sic!*] buena que el diablo, they guarantee” (*Lose* 204). Y este es un ejemplo de pregunta parcial sin inversión: “They beat the anti-Pura drums daily. Ella es prieta. Ella es fea. Ella dejó un hijo en Santo Domingo. Ella tiene otro aquí. No tiene hombre. No tiene dinero. No tiene papeles. Qué tú crees que ella busca por aquí?” (*Lose* 104).

Por otra parte, establece sutiles diferencias entre términos en principio sinónimos (según el diccionario) demostrando un conocimiento profundo de las connotaciones y ámbitos de uso de cada uno, como en este ejemplo: : “Beli might have been a puta major [*sic!*] in the cosmology of her neighbors, but a cuero she was not” (*Oscar* 103), donde obviamente *puta* y *cuero*, si bien ambos significan “prostitutas”, no ocupan el mismo lugar en la escala valórica del hablante.

De este modo, al lector se le ofrece la posibilidad de efectivamente “oír” el sociolecto auténtico de los personajes de Junot Díaz, lo cual representa, a mi modo de ver, un tipo *sui-generis* de cambio de código intra-dialectal en español, subordinado al cambio de código general entre el castellano “neutro” y el inglés, que es propio de toda su prosa.

## 6. JUEGOS LINGÜÍSTICOS, CREACIONES LÉXICAS HÍBRIDAS Y FUSIÓN DE CÓDIGOS EN LA OBRA DE JUNOT DÍAZ

La predisposición lingüística lúdica de Junot Díaz va mucho más allá de la alternancia o mezcla de códigos de diferentes tipos descrita más arriba. Por ejemplo, el autor propone ingeniosos juegos de palabras, como en el

ejemplo que sigue, donde juega con la semejanza fonética entre la última parte del nombre de Trujillo —pronunciado con la *ele* palatal de algunas partes del mundo hispano— y la palabra *lío* (en que el personaje se mete por haber irritado a este tirano), para acuñar un nuevo vocablo híbrido y sumamente expresivo (*Truji-lío*):

- (33) Two Truji-líos in one life time- what in carajo else could it be? (*Oscar* 152)

Otra notable creación léxica es *tetatorio* (¿quizás un cruce entre *tetas* y *promontorio*?) , que Díaz acuña y usa (al lado del popular *pechonalidad*, ya existente<sup>12</sup>) para ponderar, en un irresistible tono humorístico, la enormidad de los senos de una muchacha:

- (34) ...her tetas were globes so implausibly titanic they made generous souls pity their bearer and drove every straight male in the vicinity to reevaluate his sorry life. And what about that supersonic culo... a culo que jalaba más que una junta de buey, ¡Dios mío! ...Gradually, Beli began to see beyond the catchcalls and the Dios mío, asesina, and the y ese tetatorio, and the que [*sic!*] pechonalidad.... (*Oscar* 92-93)

Es más: a veces Díaz simplemente se inventa formas en español que uno no usaría, pero que suenan perfectamente idiomáticas, por así decirlo--, como en *pongándome* en vez de *poniéndome*:

- (35) This is the endgame, and instead of pulling out all the stops, instead of pongándome más chivo que un chivo, I feel sorry for myself, como un parigüayo [*sic!*] sin suerte. (*Lose* 21-22)

O, al revés, crea “verbos” ingleses<sup>13</sup> sobre adverbios españoles, como en el siguiente ejemplo:

<sup>12</sup> Agradezco al lector anónimo de la primera versión de este artículo la observación de que no se trata de una creación léxica original del autor, como yo creía, sino de un término “usado popularmente con efecto humorístico en el español peninsular”.

<sup>13</sup> La creatividad léxica de Junot Díaz en español tiene un digno parangón en su creatividad léxica en inglés, especialmente cuando introduce alusiones a la cultura popular de los *comic books* y de las películas más taquilleras. No puedo resistir la tentación de citar este ejemplo acerca de la exterminación de los taínos: “When the Euros started going Hannibal Lecter on the Tainos [*sic!*]...” (*Oscar* 244); o el siguiente ejemplo, acerca de la matanza de un joven comunista por el Capitán: “In 1977 he played mazel-tov on a fifteen-year-old boy’s throat with the heel of his Florsheim” (*Oscar* 295).

- (36) There' s a girl who keeps bumping into you. You say to her  
Pero mi amor, ya. And she says: Ya yourself. (*Lose* 211)

Efectivamente, lo que llama poderosamente la atención en la obra de Junot Díaz es la manera original en que el autor, llevando este afán lúdico a un escalón superior, funde las dos lenguas en creaciones lingüísticas híbridas, que prácticamente no se pueden asignar con claridad a la gramática de ninguno de los dos idiomas involucrados, ya que participan de ambos. Por ejemplo, cuando dice *Plátano Curtain* (*Oscar* 225) —o sea, la “Cortina de Plátanos”, calco de la famosa Cortina de Hierro— para referirse a la situación de la República Dominicana bajo la dictadura de Trujillo, semejante a la situación de los países europeos del Bloque del Este durante la Guerra Fría, ¿está creando un compuesto inglés o español?

Además, este autor viola no pocas veces algunos de los principios “clásicos” de la llamada gramática del *code-switching*, como por ejemplo la restricción del morfema libre, que prohíbe la combinación de un morfema ligado con una raíz en una lengua diferente (como en \**eat-eando*, cf. Poplack 2000); o la supuesta dificultad de cambiar de lengua en medio de un constituyente nominal o verbal, como por ejemplo entre un determinante y un nombre, entre un cuantificador y un adjetivo o entre un auxiliar verbal y el verbo principal en una forma compuesta. Y sin embargo, aquí está Junot Díaz creando impunemente vocablos que tienen una raíz española y un sufijo inglés, ya sea flexivo, como en este gerundio: *perejiling* (sobre el cual voy a volver más adelante); ya sea derivativo, como en *culocracy* y *culocrat*, por ejemplo, refiriéndose a Trujillo y a su época.

- (37) They say he [*o sea Trujillo, cuando fue asesinado*] was  
on his way for some ass that night. Who is surprised? A  
consummate culocrat to the end. (*Oscar* 154)

Y en otro lugar afirma que “...the regime was the first culocracy...” (*Oscar* 217).

Otros ejemplos podrían ser *campesina-ness*, *prieta-ness*, *cursí-ness*, características usadas para describir a diversas muchachas del barrio; o la hilarante creación léxica *no-toto-itis*, para referirse a la “virginidad crónica” de su tímido compañero de cuarto (*toto* es el nombre vulgar de la “vulva”).

- (38) Problem was, when it came to the mujeres my roommate  
was like no one on the planet. On the one hand, he had the  
worst case of no-toto-it is I'd ever seen. (*Lose* 173)

También lo vemos forjando compuestos bilingües (*hybrid o mixed loans*, como se les llama en lingüística), cuando habla de *my abuelo and his campo hands, stupid campo-jokes, my family doctora*, o combinaciones más complejas, pero no menos híbridas, como *a prime cut of culo from Colombia* (Oscar 121). Aquí están algunos de estos ejemplos contextualizados:

- (39) The nigger was without doubt the family's resident méteselo expert. (Oscar 24)
- (40) In 1937, ...while the friends of the Dominican Republic were perejiling<sup>14</sup> Haitians and Haitian-Dominicans and Haitian-looking Dominicans to death... (Oscar 215)

Y aquí van algunos ejemplos adicionales, de la misma índole: “He was becoming an alguien” (Oscar 121); “He didn’t care about nada that night” (Oscar 47); “Manny was muy bald and completely shaved his head to hide it” (Oscar 43); “Your outrageous sinvergüencería” (Lose 47); y “Her querido oldest hijo, her total consentido” (Lose 107). Por otra parte, el ejemplo que sigue ilustra el cambio dentro del núcleo del constituyente verbal:

- (41) Dude was figureando hard. Had always been a papi chulo, so of course he dove right back into the grip of his old sucias. (Lose 94)

Otro rasgo digno de mencionar es que en ciertos pasajes se puede detectar la influencia del español no solo en el léxico, sino también en la sintaxis inglesa de los protagonistas, como en los siguientes dos ejemplos, que, más bien que simples errores gramaticales debidos a una adquisición incompleta de la lengua del país de acogida, parecen responder al deseo de Junot Díaz, expresado en la entrevista citada más arriba, de mostrar cómo “no sólo el inglés moldea al español, sino también el español moldea al inglés” (al menos el de estos inmigrantes). Por ejemplo, en esta réplica (ya citada previamente, en otro contexto) se nota la falta del auxiliar *do* en la negación, formada, como en español, poniendo simplemente el adverbio de negación *no* ante el verbo<sup>15</sup>.

<sup>14</sup> “El perejil” era la prueba a la que los soldados de Trujillo sometían a los negros que vivían en la frontera con Haití, para distinguir entre los haitianos y los dominicanos, ya que los haitianos no podían pronunciar bien la *r* en palabras como *perejil*: si, cuando les mostraban la respectiva planta y les pedían que la nombraran, no lo hacían bien, los mataban (comunicación personal de Porfirio Rodríguez). Este matiz —y muchos otros, por lo demás— se pierden, como ya he dicho, en la versión española de la novela (Díaz 2008), ya que el traductor usa simplemente “perejiliando” como si este fuese un verbo común y corriente del español.

<sup>15</sup> Es uno de los casos de *code-switching* “imposibles”, comentados en Timm (1975).

(42) You no deserve I speak to you in Spanish, he says. (*Lose* 4)

Y en este pasaje se nota la elisión del pronombre sujeto *she* en la segunda oración, elisión que es normal en español, pero no está permitida en inglés (y, sin embargo, aparece con cierta frecuencia en la prosa de Junot Díaz).

(43) Yo soy prieta, Yuni, she said, pero no soy bruta. Knew exactly what kind of sucio I was. (*Lose* 69)

También sospecho que la -s final en el verbo *sacar*, en la segunda réplica de este diálogo:

(44) Find yourself another girl, Elvis advises... Clavo saca clavo. Nothing sacas nothing, you reply. No one will ever be like her. (*Lose* 81)

puede no ser un error tipográfico, sino que pretende ser la terminación verbal de los verbos ingleses en tercera persona singular, aplicada a una raíz española, de la misma manera que la s en “his tío’s car” o “the pueblo’s money” (*Oscar* 139) representa el genitivo sajón aplicado a unas voces latinas<sup>16</sup>.

Los ejemplos comentados en este apartado inducen a pensar que para Junot Díaz el español y el inglés no solo alternan en pasajes o fragmentos textuales sucesivos, como se ha visto en el apartado anterior, sino que incluso tienden, muchas veces, a fusionar el uno con el otro, en un único código discursivo cuyos hilos lingüísticos constitutivos resultan, en última instancia, si no imposibles, al menos muy difíciles de desentrañar. Y no me refiero solamente a las creaciones léxicas o a los sintagmas híbridos arriba mencionados, sino también a los casos, numerosísimos, en que la alternancia es tan constante que, a cierto punto, la frontera entre la lengua matriz y la lengua que se inserta en esta matriz (según se ha planteado en la literatura de especialidad sobre *code-switching*—cf. Meyers Scotton 1993, en particular) se vuelve borrosa<sup>17</sup>.

<sup>16</sup> De hecho, el proceso de “hispanización” del inglés en las obras literarias de autores latinos ha sido observado por varios investigadores del *code-switching*; entre ellos Aparicio (1994) —quien habla del “inglés tropicalizado” de varios escritores chicanos—, Álvarez (2013) —quien analiza el “inglés subversivo” (otro tipo de *Spanglish*, como dice ella) de un autor cubano-americano—, o Rudin (1996) —quien comenta el “inglés hispanizado” de la literatura mexicoamericana—.

<sup>17</sup> La mordaz descripción de la hermana de Trujillo ofrece un perfecto ejemplo, entre muchos otros (usados, varios de ellos, en el texto de este artículo) de integración de las dos lenguas en un mismo pasaje: “...a bitter tacaña... the time before her brother’s rise has made her un mujer bien fuerte y bien cruel; she was no pendeja and ate girls like Beli like they were pan de agua...” (*Oscar* 139).

Propongo llamar a este fenómeno “fusión de códigos”, para diferenciarlo de los casos “clásicos”, por así decirlo, de alternancia de códigos inter- o intra-oracionales.

Pero, antes de continuar, conviene recordar, a estas alturas, la clasificación de los tipos de alternancia de códigos intraoracionales (o *code-mixing*, como prefiere él llamarlos) de Muyskin (2000) – a saber la inserción, la alternancia y la lexicalización congruente<sup>18</sup>–, una clasificación que ha resultado un muy fructífero punto de partida para algunos estudios recientes del cambio de códigos en la literatura. Un ejemplo sería el excelente estudio sobre el cambio de códigos en *The Brief Wondrous Life of Oscar Wao*, por Eugenia Casielles-Suárez (2013), donde se plantea que el “bilingüismo radical” se puede conseguir no solo a través de una alternancia de códigos repetida (como es el caso de los textos analizados en Torres 2007, que introduce este concepto), sino también –aunque con efectos distintos y más potentes– a través de la frecuente inserción y lexicalización congruente. Dice Casielles-Suárez, con respecto a la novela analizada:

In addition to the extensive borrowing of single words ..., in Díaz’s novel we find all three types of code-switching. However, in contrast to other radical bilingual texts ... which make use of sustained code-alternation, and where the switch usually involves more than one constituent, what is striking about this text is the abundant use of insertion and congruent lexicalization. (2013: 483)

Su conclusión es que en el caso de Díaz hay una conexión más íntima entre las dos lenguas, ya que el español, más que alternar con el inglés, lo “invade” y se convierte en una parte constitutiva del mismo<sup>19</sup>. Por lo tanto, Casielles-Suárez propone distinguir entre *radical bilingualism* (basado en *sustained alternation*, y propio de textos como los de Braschi y Chávez-Silverman<sup>20</sup>) y lo que ella llama *radical hybridism*, refiriéndose a “the use

<sup>18</sup> El autor citado define estos conceptos de la siguiente manera: “**insertion** of material (lexical items or entire constituents) from one language into a structure from the other language; **alternation** between structures from languages; **congruent lexicalization** of material from different lexical inventories into a shared grammatical structure” (Muysken 2000:3).

<sup>19</sup> Lo mismo piensa Ch’ien cuando incluye en su ensayo sobre Díaz –uno de los representantes más destacados, para ella, del “weird English”, al lado de otras personalidades contemporáneas, como Salman Rushdie, por ejemplo– una larga lista de citas de la novela en cuestión, en las que “Spanish is an assumed American language” (2004: 205-206).

<sup>20</sup> Ver Derrick (2014) sobre el bilingüismo radical en esta autora.

of spontaneous loan words, and sustained insertion and congruent lexical à la Díaz”, que hace que “rather than alternating with English, Spanish words, hundred of Spanish words and phrases, blend in with English grammar and are treated as if they were English” (485).

Me parece que el “hibridismo radical” planteado por esta autora corresponde, en esencia, al concepto de “fusión de códigos” que yo había propuesto antes de conocer su excelente artículo (que me ayudó mucho a clarificar y entender mejor mi propia percepción del texto de Díaz). Esta fusión de códigos (que traduciría al inglés como *code-fusion*) me sirve ahora para hacer la transición hacia el fenómeno del translenguar, con el cual quiero concluir.

## 7. A GUISA DE CONCLUSIÓN: EL TRANSLENGUAR EN JUNOT DÍAZ

Para concluir, quisiera proponer que esta fusión de códigos en la obra de Junot Díaz representa un tipo de translenguar transplantado a la literatura. El translenguar (en inglés, *translanguaging*), es un concepto que Ofelia García, en un estudio reciente<sup>21</sup>, define como:

[E]l conjunto de prácticas discursivas complejas de todos los bilingües y las estrategias pedagógicas que utilizan estas prácticas discursivas para liberar las maneras de hablar, ser y conocer de comunidades bilingües subalternas”.  
(2013: 354).

El translenguar —continúa García— “es diferente del denominado ‘code-switching’ en inglés, ya que no se refiere a una alternancia o cambio de código, sino al uso de prácticas discursivas que, vistas desde una perspectiva bilingüe, no pueden ser fácilmente asignadas a una u otra lengua” (2013: 354). Para esta autora, el translenguar (que se inspira en el concepto de “trasculturación” del etnólogo cubano Fernando Ortiz y fue usado primero por Cen Williams- cf. García y Wei 2013):

<sup>21</sup> Los lectores interesados pueden encontrar información mucho más completa sobre este tema en la reciente monografía que García y Wei (2013) dedican a este importante concepto educativo. Según estos autores, al origen, el término, que viene del galés *trawsieithu*, designaba una práctica pedagógica en la que a los estudiantes se les pedía que alternaran dos lenguas durante su uso receptivo o productivo. También Sayer (2013) ejemplifica su uso por los educadores estadounidenses en el aula escolar.



[S]e construye dentro de una epistemología post-estructuralista y post-colonialista, que reconoce la heteroglosia de las prácticas discursivas de los hablantes, sobre todo los multilingües, y las características dinámicas que las definen” (2013: 354).

Los escritos de Junot Díaz cumplen fielmente, en esencia, con las dos características fundamentales del “translenguar” tales como las describen García y Wei, cuando afirman:

Translanguaging... embraces both *creativity*; that is, following or flouting norms of language use, as well as *criticality*; that is, using evidence to question, problematize or express views. (García y Wei, 2004: 27).

Como se ha comentado anteriormente, Díaz crea con la mezcla de idiomas “rompiendo las normas” del uso común y corriente (particularmente entre monolingües) y lo hace para cuestionar el *status quo* y para expresar su concepción de que el español, en Estados Unidos, no es una lengua “extranjera” subordinada al inglés, sino parte constitutiva del tejido lingüístico estadounidense en su conjunto. Es, como se expresó plásticamente años atrás Carlos Alonso (2007), “our foreign national language”. Y el Premio Pulitzer que se ha ganado Junot Díaz parece efectivamente consagrarlo como un representante destacado no del *Spanglish*, como se ha dicho a veces, sino de la práctica discursiva heteroglósica típica de nuestra sociedad global del siglo XXI, que es el translenguar. Quiero dejar al lector con esta incitante reflexión de García y Wei, acerca de las diferencias entre los dos conceptos:

The appropriation of these terms [i.e. Spanglish] by ethnolinguistic communities points to the growing fluidity in bilingual speech. But these terms take as their point of departure a ‘standard’ language, comparing bilingual speech to each of the ‘languages,’ as defined and described by the dominant members of society. In contrast, translanguaging assumes one linguistic repertoire that could never be split into one or another language, an Aleph in the Borgean sense that contains the sum total of the meaning-making universe of bilingual speakers. (García y Wei 2013: 48)





## REFERENCES

- Alonso, Carlos (2007): "Spanish: The Foreign National Language", *Profession 2007*, pp. 218-228.  
<http://dx.doi.org/10.1632/prof.2007.2007.1.218>
- Álvarez, Celia (1991): "Code-Switching in Narrative Performance: Social, Structural and Pragmatic Function in the Puerto Rican Speech Community of East Harlem", en Carol A. Klee y Luis A. Ramos García (eds.): *Sociolinguistics of the Spanish-Speaking World: Iberia, Latin America, United States*. Tempe, Arizona: Bilingual Press/Editorial Bilingüe, pp. 271-298.
- Álvarez, Stephanie (2013): "Subversive English in *Raining Backwards*: A Different Kind of Spanglish", *Hispania* 96.3, pp. 444-459.  
<http://dx.doi.org/10.1353/hpn.2013.0067>
- Anzaldúa, Gloria (1999 [1987]): *Borderlands/La frontera. The New Mestiza*. San Francisco: Aunt Lute Books.
- Aparicio, Frances (1994): "On Sub-Versive Signifiers: U.S. Latina/o Writers Tropicalize English", *American Literature* 66.4, pp. 795-801.  
<http://dx.doi.org/10.2307/2927701>
- Auer, Peter (ed.) (1998): *Code-Switching in Conversation: Language, Interaction and Identity*. London & New York: Routledge.
- Baez, Nalda R. (2008): "La experiencia del exilio: La clase social, el idioma, los géneros sexuales y la identidad en textos seleccionados de Julia Álvarez, Loida Maritza Pérez y Junot Díaz: Escritores dominico-americanos". Tesis doctoral, Purdue University.
- Basavilbaso, Teodelina (2013): "Estados Unidos les tiene fobia a los latinos" (Entrevista con Junot Díaz). *La Nación*, 11 de octubre, 2013. ([www.lanacion.com.ar](http://www.lanacion.com.ar)).

- Becker, Kristin R. (1997): "Spanish-English Bilingual Code-switching: A Syncretic Model". *Bilingual Review* 22.1, pp. 3-31.
- Betti, Silvia (2008): *El Spanglish: ¿medio eficaz de comunicación?* Bologna: Pitagora Editrice.
- Bullock, Barbara E., y Almeida Jacqueline Toribio (eds.) (2009): *The Cambridge Handbook of Linguistic Code-switching*. Cambridge: Cambridge University Press.  
<http://dx.doi.org/10.1017/CBO9780511576331>
- Bürki, Yvette (2003): "La alternancia de códigos en la literatura neorriqueña", *Revista Internacional de Lingüística Iberoamericana* 1.2, pp. 79-96.
- (2008): "La representación de la oralidad bilingüe en la literatura: *Caramelo*", en Jenny Brumme/Hildegard Resinge (eds.): *La oralidad fingida: Obras literarias. Descripción y traducción*. Madrid-Frankfurt : Iberoamericana/Vervuert, pp. 33-58.
- Burrows, Sonja S. (2010): "Beyond the Comfort Zone: Monolingual Ideologies, Bilingual U.S. Latino Texts". Tesis doctoral, Universidad de Oregon.
- Callahan, Laura (2002): "The Matrix Language Frame Model and Spanish/English Codeswitching in Fiction", *Language and Communication* 22, pp.1-16.  
[http://dx.doi.org/10.1016/S0271-5309\(01\)00018-0](http://dx.doi.org/10.1016/S0271-5309(01)00018-0)
- (2003): "The Role of Register in Spanish-English Codeswitching in Prose", *Bilingual Review* 27.1, pp. 2-25.
- (2004): *Spanish/English Codeswitching in a Written Corpus*. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- (2009): *Spanish and English in U.S. Service Encounters*. New York: Palgrave Macmillan.  
<http://dx.doi.org/10.1057/9780230619104>
- Carvalho, Ana María (2012): "Code-Switching: From Theoretical to Pedagogical Considerations", en Sara M. Beaudrie y Marta Fairclough (eds.): *Spanish as a Heritage Language in the United States: The State of the Field*. Washington DC: Georgetown University Press, pp. 139-157.
- Casielles-Suárez, Eugenia (2013): "Radical Code-switching in *The Brief Wondrous Life of Oscar Wao*", *Bulletin of Hispanic Studies* 90.4, pp. 475-487.  
<http://dx.doi.org/10.3828/bhs.2013.30>

- Ch'ien, Evelyn Nien-Ming (2004): *Weird English*. Cambridge: Harvard University Press.
- Derrick, Roshawnda (2014): "Radical Bilingualism in *Killer Crónicas*", ponencia presentada en el 129 Congreso Anual de la Asociación de Lenguas Modernas (MLA), Chicago, 10 de enero de 2014 (por aparecer en *Hispania* 97.3, en forma condensada)
- Díaz, Junot (2007): *The Brief Wondrous Life of Oscar Wao*. Nueva York: Riverhead Books.
- (2008): *La breve y maravillosa vida de Óscar Wao*. Trad. de Achy Obejas. Nueva York: Vintage Español.
  - (2012): *This Is How You Lose Her*. New York: Riverhead Books.
- Díaz Pérez, Francisco Javier (2012): "Writing and Translating on the *Frontera*: Language and Identity in Liliana Valenzuela's Translations of Latina Writers in the United States", en Catalina Iliescu Gheorghiu (ed.): *Traducción y (A)culturación en la era global/Translation and (Ac)culturation in the Global Era*. Alicante: Agua Clara, pp. 171-180.
- D'Introno, Francesco (1996): "Spanish-English Code-Switching: Conditions on Movement", en Ana Roca y John B. Jensen (eds.): *Spanish in Contact: Issues in Bilingualism*. Somerville: Cascadilla Press, pp. 187-201.
- Dumitrescu, Domnita (2010): "Spanglish: An Ongoing Controversy", en Susana Rivera-Mills y Juan Antonio Trujillo (eds.): *Building Communities and Making Connections*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 136-167.
- (2011): "Elespañol en los Estados Unidos: crecimiento, metamorfosis y controversia", *Boletín de la Academia Norteamericana de la Lengua Española* 14, pp. 261-302.
  - (2012 a): "'Spanglish': What's in a Name?" (Guest Editorial). *Hispania* 95.3, pp. ix-xii.
  - (2012 b): "The Representation of Regional Spanish Speech in Literary Dialogues from the Past Century", en Alain Létourneau, François Cooren y Nicholas Bencherki (eds.): *Representations in Dialogue/ Dialogue in Representations: Proceedings of the 13th Conference of the International Association for Dialogue Analysis (IADA)*. Montreal: John Benjamins, pp. 221-233.
  - (2012 c): "Rum. *cică* vs. esp. *dizque*: Polifonía e intertextualidad", en Clara Ubaldina Lorda Mur (coord.): *Polifonía e intertextualidad en el diálogo*. Madrid: Arco Libros, pp. 317-337.

- (2013a): “‘Two Tongues That Come Together...’ o el español en contacto con el inglés en los Estados Unidos”, *Ventana Abierta* 34, pp. 12-14.
  - (2013b): “Lo que es y lo que no es: Una nota sobre el *Spanglish*”, *Revista de la Academia Norteamericana de la Lengua Española (RANLE)*, vol.2. núm.4, pp. 353-361.
  - (2013c): “El español en la obra de Junot Díaz”, ponencia presentada en el XVIII Congreso de la Asociación Internacional de Hispanistas, Buenos Aires, 17 de julio de 2013.
  - (2014): “La búsqueda de la poesía plural y plurilingüe en Octavio Paz”, en Roberto Cantú (ed.): *The Willow and the Spiral: Essays on Octavio Paz and the Poetic Imagination*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp.233-243.
- Flores, Lauro (1987): “Converging Languages in a World of Conflicts: Code-Switching in Chicano Poetry”, *Visible Language* 21.1, pp. 130-152.
- Franceschini, Rita (1998): “Code-switching and the Notion of Code in Linguistics: Proposals for a Dual Focus Model”, en Peter Auer (ed.): *Code-Switching in Conversation: Language, Interaction and Identity*. London and New York: Routledge, pp.51-72.
- Frydman, Jason (2009): “Violence, Masculinity, and Upward Mobility in the Dominican Diaspora: Junot Díaz, the Media, and *Drown*”, en Harold Bloom (ed.): *Hispanic-American Writers* (New Edition). New York: Infobase Publishing /Bloom’s Literary Criticism, pp. 133-143.
- Fuller, Janet M. (2013): *Spanish Speakers in the USA*. Bristol/Buffalo/Toronto: Multilingual Matters.
- García, Mary Ellen. (2009): “Code-Switching and Discourse Style in a Chicano Community”, en Manel Lacorte y Jennifer Leeman (eds.): *Español en Estados Unidos y otros contextos de contacto: Sociolingüística, ideología y pedagogía/ Spanish in the United States and Other Contact Environments: Sociolinguistics, Ideology, and Pedagogy*. Madrid/Frankfurt: Iberoamericana/Vervuert, pp. 137-156.
- García, Ofelia (2013): “El papel del *translenguar* en la enseñanza del español en los Estados Unidos”, en Domnita Dumitrescu y Gerardo Piña-Rosales (eds.): *El español en los Estados Unidos: E Pluribus Unum? Enfoques multidisciplinares*. New York: ANLE, pp. 353-373.

- y Li Wei (2014): *Translanguaging: Language, Bilingualism and Education*. London: Palgrave MacMillan.
- García Vizcaíno, María José (2008): “Cisneros’ Code-Mixed Narrative and Its Implications for Translation”, *Mutatis Mutandis* 1.2, pp. 212-224.
- Gimeno Menéndez, Francisco y María Victoria Gimeno Menéndez (2003): *El desplazamiento lingüístico del español por el inglés*. Madrid: Cátedra.
- Hansen Esplin, Marlene (2012): “Spanish, English, and in-between: Self-Translation in the U.S. and Latin America”. Tesis doctoral, Michigan State University.
- Heller, Monica (ed.) (1988): *Codeswitching: Anthropological and Sociolinguistic Perspectives*. Berlin: Mouton de Gruyter.  
<http://dx.doi.org/10.1515/9783110849615>
- Jonsson, Carla (2010): “Functions of Code-Switching in Bilingual Theater: An Analysis of Three Chicano Plays”, *Journal of Pragmatics* 42, pp. 1296-1310.  
<http://dx.doi.org/10.1016/j.pragma.2009.08.012>
- Kakutani, Michiko (2007): “Travails of an Outcast”. *The New York Times*, 4 de septiembre de 2007.
- Kecskes, Istvan (2006): “The Dual Language Model to Explain Code-Switching: A Cognitive-Pragmatic Approach”, *Intercultural Pragmatics* 3.3, pp. 257- 283.  
<http://dx.doi.org/10.1515/IP.2006.017>
- Keller, Garry D. (1984). “How Chicano Authors Use Bilingual Techniques for Literary Effect”, en Isidro Ortiz, Eugene García y Francisco Lomelí. (eds.): *Chicano Studies: A Multidisciplinary Approach*. New York: Teachers College, pp. 171-192.
- y Randall G. Keller (1993): “The Literary Language of United States Hispanics”. In Francisco Lomelí (ed.): *Handbook of Hispanic Cultures in the United States: Literature and Art*. University of Houston: Arte Público Press, pp. 163-183.
- Kevane, Bridget (2003): *Latino Literature in America*. Westport, CT: Greenwood Press.
- Klee, Carol A. y Andrew Lynch (2009): *El español en contacto con otras lenguas*. Washington D.C.: Georgetown University Press.
- Koike, Dale April (1987): “Code-Switching in the Bilingual Chicano Narrative”, *Hispania* 70.1, pp. 148-154.  
<http://dx.doi.org/10.2307/343692>

- Ladouceur, Louise (2006): "Write to Speak: Accents et alternances de codes dans les textes dramatiques écrits et traduits au Canada", *Target* 18.1, pp. 49-68.  
<http://dx.doi.org/10.1075/target.18.1.04lad>
- Lago, Eduardo (2013). "Cada joven es un objetivo de las corporaciones" (Entrevista con Junot Díaz). *El País Semanal*, 29 de abril de 2013.
- Lipski, John (1982): "Spanish-English Language Switching in Speech and Literature: Theories and Models", *The Bilingual Review/La Revista Bilingüe* 9.3, pp.191-212.
- (2008): *Varieties of Spanish in the United States*. Washington D.C: Georgetown University Press.
- López Morales, Humberto (coord.) (2009): *Enciclopedia del español en los Estados Unidos* (Anuario del Instituto Cervantes 2008). Madrid: Cervantes /Santillana.
- Mendieta-Lombardo, Eva y Zaida A. Cintron (1995): "Marked and Unmarked Choices of Code-Switching in Bilingual Poetry", *Hispania* 78.5, pp. 565-572.  
<http://dx.doi.org/10.2307/345306>
- Montes-Alcalá, Cecilia (2000): "Two Languages, One Pen: Socio-Pragmatic Functions in Written Spanish-English Code-Switching". Tesis doctoral, Universidad de California, Santa Barbara.
- (2012): "Code-Switching in U.S. Latino Novels", en Mark Sebba, Shahrzad Mahootian and Carla Jonsson (eds.): *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*. London: Routledge, pp. 68-88.
- Muysken, Pieter (2000): *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol (1993): *Duelling Languages*. Oxford: Clarendon.
- Nuessel, Frank (2000): *Linguistic Approaches to Hispanic Literature*. Ottawa: Legas.
- Otheguy, Ricardo (2013): "Convergencia conceptual y la sobrestimación de la presencia de elementos estructurales ingleses en el español estadounidense", en Domnita Dumitrescu y Gerardo Piña-Rosales (eds.): *El español en los Estados Unidos: E Pluribus Unum? Enfoques multidisciplinarios*. New York: ANLE, pp. 129-149.
- y Nancy Stern (2010): "On So-called Spanglish", *International Journal of Bilingualism* 15.1, pp.85-100.  
<http://dx.doi.org/10.1177/1367006910379298>



- Pfaff, C., and L. Chávez (1986): "Spanish/ English Codeswitching: Literary Reflections of Natural Discourse", en Renate Von Bardeleben, Dietrich Briesemeister y Juan Bruce-Novoa (eds.): *Missions in Conflict: Essays on US-Mexican Relations and Chicano Culture*. Tübingen: Narr, pp. 229-254.
- Poplack, Shana. (2000/[1979/80]): "Sometimes I'll Start a Sentence in Spanish *y termino en español*: Toward a Typology of Code-Switching", en Li Wei (ed.): *The Bilingual Reader*. London & New York: Routledge, pp. 221-256.
- Rodríguez, Richard (1982): *Hunger of Memory: The Education of Richard Rodriguez (An Autobiography)*. New York: The Dial Press.
- Rodríguez-Gonzales, Eva y M. Carmen Parafita-Couto (2012): "Calling for Interdisciplinary Approaches to the Study of Spanglish and Its Linguistic Manifestations", *Hispania* 95.3, pp. 461-480.
- Rudin, Ernest (1996): *Tender Accents of Sound: Spanish in the Chicano Novel in English*. Tempe, Arizona: Bilingual Press/Editorial Bilingüe.
- Sayer, Peter (2013): "Translanguaging, TexMex, and Bilingual Pedagogy: Emergent Bilinguals Learning Through the Vernacular", *Tesol Quarterly*, 47.1, pp. 63-88.  
<http://dx.doi.org/10.1002/tesq.53>
- Silva-Corvalán, Carmen (1983): "Code-shifting Patterns in Chicano Spanish", en Lucía Elías-Olivares y D.Nasjleti (eds.): *Spanish in the United States: Beyond the Southwest*. Washington D.C.: National Center for Bilingual Education, pp. 69-87.
- (1989): *Sociolingüística: Teoría y análisis*. Madrid: Alhambra.
  - (1994): *Language Contact and Change: Spanish in Los Angeles*. Oxford: Clarendon.
  - (2001): *Sociolingüística y pragmática del español* (con Ejercicios de Reflexión de Andrés Enrique-Arias). Washington DC: Georgetown University Press.
- Silva Rodríguez, Graciela y Manuel de Jesús Hernández-G. (eds.) (2012): *Chican@s y mexican@s norteñ@s: Bi-Borderlands Dialogues on Literary and Cultural Production*. México: Eón/Arizona State.
- Timm, Lenora A. (1975): "Spanish-English Code Switching: El Porqué y How Not To", *Romance Philology* 28.4, pp. 473-482.




- (2000): “Y Se Hinchó Into Armor: The Pragmatics, Metapragmatics, and Aesthetics of Spanish/English Code-Switching Poetry”, *Southwest Journal of Linguistics* 19.2, pp. 91-114.
- Toribio, Almeida Jacqueline (2000): “Once upon a time en un lugar muy lejano...’ Spanish-English Code-Switching across Fairy Tales Narratives”, en Ana Roca (ed.): *Research on Spanish in the United States: Linguistic Issues and Challenges*. Somerville, MA: Cascadilla Press, pp. 184-203.
- y Edward J. Rubin (1996): “Code-Switching in Generative Grammar”, en Ana Roca y John B. Jensen (eds.): *Spanish in Contact: Issues in Bilingualism*. Somerville: Cascadilla Press, pp. 203-226.
- Torres, Lourdes (2007): “In the Contact Zone: Code-Switching Strategies by Latino/a Writers”, *MELUS* 32.1, pp. 75-96.  
<http://dx.doi.org/10.1093/melus/32.1.75>
- Valdés Fallís, Guadalupe (1976a): “Social Interaction and Code-Switching Patterns: A Case Study of Spanish-English Alternation”, en: Garry Keller, Richard Teschner y Silvia Viera (eds.): *Bilingualism in the Bicentennial and Beyond*. New York: Bilingual Press, pp. 86-96.
- (1976b): “Code-Switching in Bilingual Chicano Poetry”, *Hispania* 59.4, pp. 877- 886.  
<http://dx.doi.org/10.2307/340213>
- (1977): “The Sociolinguistics of Chicano Literature: Towards an Analysis of the Role and Function of Language Alternation in Contemporary Bilingual Poetry”, *Point of Contact/ Punto de Contacto* 1.4, pp. 30-39.
- Zentella, Ana Celia (1982): “Spanish and English in Contact in the United States: The Puerto Rican Experience”, *Word* 33.1-2, pp. 41-57.
- (1997): *Growing Up Bilingual: Puerto Rican Children in New York*. Malden, MA: Blackwell Publishers.

# **III**

## **Language Acquisition**





## **Heritage Language Learners: Incomplete Acquisition of Grammar in Early Childhood**

Jürgen M. Meisel  
*University of Hamburg and  
University of Calgary*

**Abstract:** The notion of ‘heritage language’ commonly refers to minority languages in bilingual settings, where they are primarily used in family-related domains. This functional restriction is said to lead to structural reduction, resulting from attrition or incomplete acquisition. From a developmental perspective, heritage language learners constitute a heterogeneous group including bilingual L1 and child and adult L2 learners. I argue that heritage language studies can only make meaningful contributions to acquisition research if they differentiate between attrition and incomplete acquisition and between learner types. Incompleteness suggests that the attained knowledge differs qualitatively from native systems, but the term is also used to refer to different yet not incomplete grammars. For acquisition theory, it is crucial to determine the conditions under which children exposed to a language from birth fail to develop native grammars. Heritage language studies claim that insufficient exposure to a language causes this to happen. However, they do not provide adequate information about the required quantity or quality of input.

**Keywords:** Heritage languages, incomplete acquisition, input, acquisition types.



## 1. INTRODUCTION

Every natural language might, in principle, qualify as a 'heritage language' (HL) if children are acculturated in it; however, the term commonly refers to bilingual communities where, in addition to the majority language of the country of residence, a non-autochthonous minority language is spoken. This is typically the language of origin of immigrants, and its use is frequently confined to family-related sociolinguistic domains. The notion of HL is thus primarily a socio-political one which translates into sociolinguistic definitions of linguistic communities.

The main defining criteria are the national origin of the families, the place of birth of HL speakers, their cultural self-identification, their age at the time of immigration (if not born in the country of residence), languages spoken in the family, languages spoken with siblings, formal education in the HL, frequency of use of the HL, whether the HL is used actively and in which social contexts, and proficiency (self-rated or tested) in the HL; cf. Alarcón (2011), Foote (2010), Sekerina and Trueswell (2011), to mention only some recent studies.

The interest recently shown in heritage languages stems from the fact that they are commonly analyzed as reduced or simplified varieties of the respective national languages; cf. Montrul (2008), Polinsky (2006), Rothman (2009), Silva-Corvalán (2003), among others. In other words, they are understood as representing a cultural heritage in just those situations where this heritage is in danger of being lost. It is this process of gradual decline which Silva-Corvalán (2003: 375) refers to in the following quote.

(I)n situations of societal bilingualism, the functionally restricted language evidences grammatical simplification, among other phenomena. The question that arises is whether this stage of grammatical simplification is due to imperfect or incomplete acquisition in the early years of life, or a result of processes of attrition or loss of acquired knowledge of the underused language.

In fact, it has frequently been observed that not all domains of grammar are likely to be reduced or simplified. Consequently, another important issue concerns the question of which aspects of grammar are affected by such changes; see Rothman (2009: 158). Moreover, to the extent that

these changes indeed happen in the course of acquisition or in subsequent attrition processes, one also needs to ask what causes incomplete acquisition or attrition to happen. Impoverished input is clearly the most obvious and frequently suggested factor, but in view of the fact that limited input does not necessarily lead to incomplete acquisition, as has been shown in research on child bilingualism (Meisel 2007b), the answer to this question is not as straightforward as it may appear to be.

We are, however, not without relevant research to guide us. The literature on the *Pidginization Hypothesis*, which argues that pidgins and creoles are shaped by similar processes and mechanisms as first and second language acquisition, developed similar ideas in quite some detail; see, for example, the contributions to the volume edited by Andersen, especially the introduction by Andersen (1983). The suggestion by Schlyter (1993) that the *weaker* language of bilinguals shares crucial properties with second languages is also directly relevant to debates about HL speakers; cf. Meisel (2007a) for a critical discussion of this claim.

In the present contribution, I adopt the research agenda suggested by Silva-Corvalán (2003), according to which we should investigate whether grammatical characteristics of HL's result from imperfect or incomplete acquisition during early childhood or from attrition or loss of acquired knowledge. Although studies of HL speakers tend to refer indistinguishably to "attrition or incomplete acquisition", these two types of processes must indeed be distinguished, for even in cases where they result in similar kinds of structural reductions of the HL, cf. Polinsky (2011), they differ significantly in terms of the cognitive processes involved. In what follows, I will, nevertheless, focus on the acquisition of heritage languages and, more specifically, on alleged cases of incomplete acquisition. In doing so, I adopt the theoretical framework of Universal Grammar (UG), as do most of the HL studies referred to in this paper.

## 2. HERITAGE LANGUAGE LEARNERS

From the perspective of language acquisition research, the first observation is that heritage language *learners* (HLL), much like heritage *languages*, are primarily defined in sociolinguistic terms; see Rothman (2009). Consequently, they represent an extremely heterogeneous group of learners. In fact, virtually all acquisition types have been subsumed

under the label of HLL: bilingual (simultaneous, 2L1) first language development, successive acquisition of languages in childhood (cL2) and in adulthood (L2), relearning after attrition or loss, to mention only the most important ones.

It goes without saying that this is not a satisfactory situation. This is not merely a terminological issue. Studies of HLL can only make meaningful contributions to our understanding of the human Language Making Capacity (LMC) and of the processes underlying language acquisition – and *vice versa*, acquisition research can only shed light on the nature and genesis of HL – if HLL are differentiated according to the criteria which serve to distinguish acquisition types, most importantly age of onset of acquisition (AO), but also type of learning environment, exposure to the primary linguistic data, etc. Differentiation of HL speakers in this fashion is a necessary prerequisite in order to be able to determine how HLL compare to known learner types with respect to processes underlying linguistic development and structural properties attested in language use. The logic of this argument also calls for a clear distinction between partially successful acquisition and attrition, notwithstanding the possibility that attrition might superficially resemble cases of acquisition which do not attain full native competence.

The question thus is whether HLL represent a specific type of learner or rather whether at least some of those referred to as HLL are unique in this sense. What transpires from the literature is that HL speakers are claimed to resemble either simultaneous or early successive bilinguals who differ, however, from both native and aL2 speakers, cf. Albirini, Benmamoun and Saadah (2011: 274). As opposed to L2 learners, such HLL were exposed to the target language from birth, but their exposure to the HL changed significantly in amount and quality during the first five or six years of life, i.e. during an age period when grammatical development in the HL was not yet completed. In other words, we are looking here at very heterogeneous learner groups who are likely to correspond to different learner types. Moreover, contrasting HLL to first language and to adult second language learners does not exhaust the possibilities of comparison. This will become clearer if we briefly recall what is currently known about bilingual acquisition. For these reasons, I believe that the conclusion is premature that we are dealing here with one acquisition type, specific of HL speakers.



### 3. SIMULTANEOUS AND SUCCESSIVE ACQUISITION OF CHILD BILINGUALISM: DEVELOPMENTAL CHANGES IN THE LANGUAGE MAKING CAPACITY

In order to be able to compare HLL to simultaneous and successive bilinguals, let me give a quick summary of some of the major claims and assumptions concerning these acquisition types. My focus is entirely on the acquisition of grammar, more specifically of morpho-syntax. By comparing different types of language acquisition, we will gain insights into the nature of the human Language Making Capacity (Slobin 1985). The LMC is assumed to be subserved by species-specific biological and cognitive mechanisms as well as by mechanisms shared with other species. The LMC becomes available to the individual as a result of neural maturation, provided it is activated during specific developmental periods by exposure to language in communicative settings. The study of only partially successful HLL might reveal some of the limits of the LMC and especially of the Language Acquisition Device (LAD, cf. McNeill 1966, Chomsky 1981). As opposed to the LMC, the LAD is understood as assembling only domain-specific principles and mechanisms, exclusively subserving the acquisition and processing of formal properties of human languages; see Meisel (2011).

Children acquiring two (or more) language simultaneously from birth are able to develop two (and probably more) native competences. If, however, the languages are acquired successively, age of onset of acquisition plays a decisive role. If AO happens at around age 10 – 12 or later, the course of acquisition and the ultimately attained grammatical knowledge differ substantively from monolingual or bilingual first language acquisition. The fact that (adult) second language acquisition (aL2) differs from (2)L1 acquisition is not controversial. Controversy exists, however, about the nature of these differences, more specifically whether they are of a fundamental nature, reflecting qualitative differences in the underlying grammatical knowledge. In the generative paradigm, this amounts to questioning whether the linguistic knowledge of L2 learners is fully constrained by principles of UG. What is also controversial is whether none or almost none (2 – 3 %) of aL2 learners attain a native-like L2 competence.

In the present context, I will refrain from engaging in a discussion of these controversial issues. It must suffice to acknowledge the existence of substantive differences between L1 and aL2. As for the question of whether or not these should be explained in terms of UG accessibility on the part of L2 learners, I can only conclude that the claim that adult L2 learners have

full access to UG has never been substantiated by empirical evidence; cf. Meisel (2011). Rather the available evidence supports not only the idea that fundamental differences between L1 and L2 acquisition exist but also that this is due to the fact that adult L2 learners have only partial access to UG, i.e. merely to invariant but not to parameterized UG principles.

These age-related changes in the LMC have been explained as being primarily caused by effects of brain maturation, this is to say, the ability to develop a full grammatical competence by mere exposure weakens or disappears in the course of development. But, importantly, it is not all of 'language' which is affected by maturational changes but certain *domains of grammar*. Moreover, the domains subject to maturational changes are not all affected simultaneously. Not only have phonology, morphology and syntax been demonstrated to follow distinct developmental agendas, we may assume that development is asynchronous even within these subcomponents of grammar. Consequently, maturational changes result not in one critical period for language acquisition. Rather, grammatical development is characterized by a number of *sensitive phases*, some of which *cluster* during various developmental phases. Each of these clusters can be regarded as constituting the equivalent of one of *multiple critical periods*; cf. Seliger (1978), Schachter (1996), Meisel (2013).

From these considerations, it follows that the LMC and its components change continuously. Critical periods or sensitive phases for language acquisition should therefore not be seen as dramatically exceptional periods in the course of a child's development but as phases of increased developmental activity during which the LAD undergoes changes which are of particular relevance for the acquisition of grammatical competence. This scenario is in accordance with the conclusion to which Hyltenstam and Abrahamsson (2003) arrive in their comprehensive state-of-the-art survey, that maturational effects on language development are noticeable as of birth and up to approximately age 15. Concerning the decisive age ranges for grammatical development, linguistic as well as neuropsychological research suggests that the age at around 7 is one during which significant changes happen, but recent findings also identify the age at around 3;6 as a critical age period affecting morpho-syntax. Critical period effects should thus be detectable between the ages of approximately 3;6 and 15 years, with the age range between approximately 4 and 6 years being of exceptional importance for morphosyntactic development. As stated by Hyltenstam and Abrahamsson (2003: 575), up to that age, all or most learners "will reach levels that allow

them to pass as native speakers - provided that there is sufficient input and that the learning circumstances are not deficient.” After this age, social-psychological factors play an increasingly important role in L2 acquisition. For a more detailed discussion of these issues, see Meisel (2013).

Supporting evidence concerning the age ranges during which maturational changes affecting grammatical development are likely to happen is provided by the neurosciences. The basic assumption is that changes in the functional organization of the brain in language processing should result in different activation patterns and in a different spatial organization of the brain if the onset of exposure to a language falls outside the optimal period, whereas no such differences should emerge if first exposure falls within this period. Neuroimaging studies indeed support the hypothesis that age of onset of acquisition is a major cause for L1-L2 differences in processing grammatical information. They furthermore corroborate the claim that important changes happen at around age 6-7, and some ERP results confirm the assumption that crucial changes occur already at around age 4; cf. Meisel (2013) for more details.

There is one further aspect concerning maturational changes of the LMC which must be mentioned here. Although the period from approximately 3-4 through 6-7 years of age has been argued to be the crucial one for the acquisition of native grammatical competence, age effects are not limited to these sensitive phases. Rather, once acquired, grammatical knowledge apparently needs to be consolidated. Studies of language attrition in individuals, e.g. Flores (2008) or Bylund (2009), have shown that “grammatical knowledge is more likely to suffer from attrition if the speaker loses contact with a language before the age of eleven” (Flores 2010: 545). This period of stabilization up to age 11-12 ensures that the knowledge remains permanently accessible; without sufficient exposure to the language during this age range, an individual’s ability to control the acquired knowledge is likely to dwindle.

Returning to the possible consequences for successive language acquisition, recent research findings suggest that if AO happens during the fourth year of life or later, acquisition still resembles L1 development in many respects, but it already exhibits aL2 properties in some domains of grammar. Successive acquisition of languages in early childhood (AO ranging approximately from 3;6 through 6 – 7) can thus be referred to as child L2 acquisition. There is still uncertainty and disagreement, on which domains of grammar are most likely to resemble aL2, but those for which empirical evidence has been provided include finiteness, tense and aspect,

root infinitives, gender assignment and agreement, clitic status of subject and object clitics and placement of object clitics if the target requires placement in a position which differs from that of the respective DP. This primarily concerns inflectional morphology, but also word order, especially if the target is not SVO (V2, OV/VO).

In sum, there can be little doubt that exposure to primary linguistic data (PLD) from the target language during sensitive phases is a necessary condition for the development of a native competence. But it is not a sufficient condition, for it presupposes adequate exposure to the PLD during periods of heightened sensitivity. It should also be noted that the existence of multiple sensitive phases implies that successful acquisition is not an all-or-nothing option, for the available input may have been sufficient for successful acquisition of some but not all aspects of grammar. In other words, we are dealing with a *three-dimensional* variation space, referring to particular grammatical features, to quantity and quality of input, and to the developmental phase during which windows of opportunity open and close. Consequently, we should expect to find, across HL learners and within individuals, that some of the vulnerable domains of grammar are fully developed whereas others are only partially acquired. We are able to identify some though not all of the grammatical domains subject to maturational changes, and we have at least an approximate idea of the crucial developmental phases, but we know virtually nothing about what might qualify as sufficient exposure. This is the issue to which research on HLL could make a significant contribution, and this is therefore the topic on which I will focus here; see below, section 4.2. First, however, we need to ask how to interpret the notion of ‘incomplete’ acquisition in terms of a learner’s developing grammar.

#### 4. INCOMPLETE ACQUISITION

The one aspect which strongly dominates the discussion of heritage languages in the linguistic literature is the fact that they are analyzed as ‘reduced’ systems. Although this is certainly not true for all HLL, the only partial success of early successive and possibly also simultaneous HL bilinguals is indeed an intriguing finding which calls for an explanation since, as we have seen, it is very well possible to acquire two first languages simultaneously. If *incomplete* acquisition is really caused by significantly reduced exposure to one of the languages during the early acquisition

phases, studies of HLL might reveal what the minimal amount of exposure to a language has to be, for learners to be able to attain a native competence. The first issue to be addressed, however, concerns the structural and psycholinguistic implications of the claim that the acquisition of a language has not been completed.

#### 4.1. On grammatical incompleteness

Most studies on HLL are rather vague as to how exactly incomplete acquisition is to be defined in structural terms and how this is reflected in mental representations of grammatical knowledge. Polinsky, for example, states simply that “incomplete acquisition... is a specific case whereby an individual fails to learn the entire system of a given language”. She claims nevertheless that the “incompletely acquired language is still a system...” (Polinsky 2006: 194). Montrul (2008: 20), on the other hand, refers to the *Incompleteness Hypothesis*, originally proposed for L2 acquisition. According to this hypothesis, incompleteness reflects representational deficits in the grammatical systems of learners. In fact, Montrul (2008: 19) explicitly states that “non-nativelike attainment is what I term in this book incomplete acquisition”. But she insists that “incompletely acquired L1 grammars diverge from the target because they are simplified, but are nonetheless constrained by Universal Grammar” (Montrul 2008: 250).

If I understand these statements correctly, they reflect different notions of grammatical incompleteness, referring implicitly or explicitly to different defining criteria. This is perhaps to be expected, and if we were to compare these uses of the term ‘incompleteness’ with those of other authors, we would undoubtedly encounter further meanings attributed to this term. But this is not about subtleties in a terminological quarrel. What matters is that we must distinguish between different kinds of achievements attained by HL learners. Only then may we expect to gain insights enhancing our understanding of language acquisition in general and of the acquisition of heritage languages in particular. This is why it is necessary to disentangle the various uses of ‘incompleteness’ which actually refer to different facts concerning the grammatical knowledge acquired by HL learners.

Let us first consider the definition according to which “an individual fails to learn the entire system of a given language”. As it stands, this statement is undoubtedly correct, but it applies to all language learners, not only to HL speakers, for all natural languages to which we refer by convenient cover terms like ‘English’ or ‘Russian’ actually comprise a number of varieties, most

importantly regional (dialects), social (sociolects) and situational (registers) ones. No individual is exposed to all the varieties which may be subsumed under the label of a given language. Consequently, no speaker of a language can be said to have acquired “the entire system”. What may be special about HL speakers – and this is probably what this definition is aiming at – is that HL learners are likely to be exposed to a more limited array of varieties, e.g. if the use of one of their languages is restricted to settings of communication within the family or among friends. This is certainly an important aspect of a sociolinguistic characterization of HL speakers which deserves to be investigated in more detail. Yet it does not provide us with insights about these speakers’ achievements as language learners, for incompleteness of grammatical knowledge results, in this case, from a lack of exposure to the relevant primary linguistic data rather than from deficient acquisition processes.

Pires and Rothman (2009) discuss this issue in some detail, distinguishing between “true incomplete acquisition” and “missing-input competence divergence”, the latter meaning that properties which are part of the competence of educated speakers are not acquired because the HL learners did not have (sufficient) exposure to formal varieties of the standard language. Pires and Rothman (2009) studied the use of inflected infinitives, a phenomenon of Portuguese grammar virtually non-existent in colloquial Brazilian Portuguese (BP) and which is therefore not acquired by speakers of BP before school age. They found that the same is true of heritage learners of BP who were born and raised in the United States, whereas US-born HL speakers of European Portuguese (EP) “have full syntactic and semantic competence of inflected infinitives”. (Pires and Rothman 2009: 230). These authors argue convincingly that the differences between EP and BP HL speakers’ knowledge of properties of inflected infinitives are due to the fact that only HL learners of EP are sufficiently exposed to primary linguistic data instantiating these properties. In view of these findings, Pires and Rothman (2009: 234) conclude that “the label incomplete acquisition is inadvertently misleading and, at worst, not explanatorily adequate”.

In accordance with this line of argument, I will use the term ‘incomplete acquisition’ to refer exclusively to cases where a phenomenon is not acquired although it is present in the ambient language of learners, because only in these situations does incompleteness reveal properties of the acquisition process rather than of the languages to which learners are exposed or of the particular settings in which acquisition happens.



The second issue to be addressed here concerns the claim that incomplete acquisition leads to non-native grammatical knowledge, as stated by Montrul (2008) who nevertheless insists that such grammars are simplified but constrained by Universal Grammar. Note that the latter claim introduces a separate and largely independent issue, one which is much too broad to be discussed in the present context; see section 3. Moreover, since it is possible to agree that HL learners acquire grammatical knowledge resembling that of L2 learners, independently of whether one agrees on whether this knowledge is fully constrained by UG, it will suffice to indicate whether or not one follows an approach implying UG conformity, thus stating explicitly what kind of knowledge is attributed to HL learners. Whether this position is well-founded, however, does not depend specifically on analyses of HL speakers, but on findings about non-native grammars more generally.

Interestingly, the approach taken by Montrul (2008) follows the Incompleteness Hypothesis, proposed by Schachter (1989, 1990, 1996) to account for differences between first and second language acquisition. According to this hypothesis, these differences are fundamental in nature, since L2 acquisition is characterized by the lack of some grammatical properties in the mental grammars of L2 learners as compared to grammars of native speakers. Crucially, this is not merely a case of grammatical divergence. Rather, incomplete grammars exhibit representational deficits, given that they lack properties which are hypothesized to be obligatory components of grammars of natural languages.

To avoid misunderstandings, a word of clarification may be useful at this point. If a grammatical notion is not implemented in an L2 linguistic system, this does not necessarily mean that the surface devices which encode the grammatical property in question are not learnable for L2 learners. But in those cases where the instantiation of this property in L1 grammars is guided by UG, e.g. by parameterized principles, L2 learners must resort to other cognitive resources in order to compensate the loss of guidance by UG. Consequently, the L2 knowledge about the target grammar can be qualified as a *hybrid system*, consisting of an incomplete version of the target grammar and a set of additional, compensatory principles which are not constrained by UG principles; see Meisel (2011). The surface expressions can thus be learned, but the linguistic knowledge of L2 learners differs qualitatively from that of native speakers. One effect of this underlying difference is that the L2 use of the properties in question is less consistent and more variable than what one typically finds in L1 speech. This observation is in line with

the finding by Sorace (2000) that optionality can persist in L2 whereas it gradually disappears in L1. She argues that the steady state of L2 acquisition tends to be extremely variable and that even advanced L2 grammars exhibit ‘residual optionality’ (Sorace 2005). From these considerations it follows that incomplete acquisition implies that the acquired knowledge is only partially a native grammar. Importantly, in order to decide whether this is indeed the case, it is not sufficient to rely on the presence of certain surface constructions.

This latter point is further corroborated by the observation that not only does the occurrence of certain surface devices not imply that the respective native grammatical notion is implemented in the non-native grammar, but, conversely, the lack or substitution of surface devices does not necessarily reflect incomplete acquisition of the underlying grammatical notion. Examples of this type of phenomena are the replacement of case endings by prepositional constructions, or the substitution of synthetic forms by analytical ones. In other words, ‘incomplete’ acquisition here refers to cases where different, possibly ‘simplified’, but qualitatively non-distinct grammars are acquired. Typological research tells us that some languages draw on pre- or postpositions where others rely on case-marking for the expression of certain grammatical functions or spatial notions. In some theories of grammar, these are treated as distinct surface realizations of the same grammatical subsystem. From this perspective, learners who choose to spell out a function using prepositions rather than case-markings can be seen to be bunting rather than hitting balls to the outfield. They are still playing baseball by the rules.

The point which I am trying to make here is that merely listing surface properties encountered in the speech of HL speakers does not allow one to assess the nature of the grammatical knowledge attained by what is qualified as ‘incomplete acquisition’. Yet many HL studies use the term ‘incomplete’ in just this fashion, referring to instances where the HL is lacking certain surface properties of the target language without thoroughly examining whether the grammatical notion encoded by the missing device is indeed not implemented in the grammar.

This is the case, for example, when Montrul (2008: 255) states that “children and adults make many of the same omission, substitution, and misplacement errors in phonology, morphology, syntax and semantics attested in L1 acquisition, all of which are a normal part of the developmental process.” She thus motivates a further extension of the already insufficiently



restrictive definition of the notion of ‘incomplete’ acquisition, referring to all types of developing grammars, monolingual as well as bilingual, as instances of “temporary” incomplete acquisition. The justification of this terminological choice is that “errors of omission, substitution, or misplacement of some obligatory grammatical element, together with restricted use of some structures, are indicative of incomplete knowledge in developing child grammars...” (Montrul 2008: 253f.). But this observation does not enhance our understanding of the nature of the grammatical knowledge attained by the various types of learners. Rather, it is more likely to obscure the picture than to shed light on it, for the claim that the developing grammars are identical is arguably wrong and certainly misleading, even if some of the surface reflections are “the same”, for they need not and, in fact, frequently do not reflect the same type of grammatical knowledge; see Meisel (2011).

There is, in fact, ample evidence provided by research publications of the 1970’s and 1980’s demonstrating that simplification processes frequently result in surface similarities in spite of fundamental knowledge differences; see Andersen (1983). Even mature native speakers use “the same” omissions, substitutions and replacements in simplifying registers like foreigner talk (*du fahren Hamburg, mit Zug, morgen?* ‘you go+inf H. with train tomorrow?’), baby talk, and so forth. But one would surely not want to claim that their grammars are incomplete when they do so; cf. Meisel (1977). These and similar issues have been studied in quite some detail in connection with the Pidginization Hypothesis, as mentioned above and in the Introduction to the present paper. The insights gained by this kind of research corroborate the claim that the surface similarities observed in the speech of different types of child and adult language learners do not reflect equivalent kinds of grammatical knowledge. Consequently, I will not adopt this notion of ‘temporary incompleteness’ but adhere to an interpretation of the term ‘incompleteness’ which implies non-native attainment of one of the grammatical systems by bilingual HLL, as originally suggested by Montrul (2008: 19).

The issue at stake here is thus to determine whether the grammars of HL learners are indeed non-native grammars and therefore ‘incomplete’ as defined by the Incompleteness Hypothesis. Implicitly, this raises broader questions as to the nature of the resulting knowledge systems, i.e. whether they are fundamentally distinct or merely divergent as compared to the systems underlying the “full languages”, an alternative described by Sorace (1993) who distinguished between incomplete and divergent representations

of grammatical knowledge. In accordance with the approach which I briefly outlined above, HL learners might acquire a kind of L2 knowledge which has been characterized as a hybrid system. But it is important to note that the assumption that HL learners develop non-native grammars does not necessarily imply that these grammars are of the same nature as those of child or adult L2 learners. To my knowledge, HL studies have not yet provided sufficient evidence in order to decide on this issue nor to answer the more general question relating to the nature of the grammatical knowledge of HL speakers.

It should be obvious by now that the characterization of HL grammars as ‘simplified’ does not offer crucial insights into the nature of the underlying system – nor does the assertion that the simplified language “is still a system” (Polinsky 2006: 194). For example, in cases where the term simplification refers to the omission of elements, this may either consist of a reduction of a redundant set of surface devices or to the loss of functionally relevant elements. In the former case, the underlying system is likely to be unaffected by the simplification. In the latter, the remaining devices must all undergo functional changes in order to fill the gap opened by the ones lost. Consider as an example a Case language with three overt cases where one is dropped. When this happens, one or both of the remaining cases may serve to express partly or fully the functions of the one which has been dropped. If this happens, the resulting system is, in fact, not incomplete, but it has undergone restructuring. Changes of this sort are well documented in diachronic change or in creolization. In the course of change from Latin to medieval Romance languages, for example, the latter retained only two morphologically marked Cases (*casus rectus*, *casus obliquus*), and they used prepositions to encode grammatical relationships which in Latin had been encoded by means of morphological case markings. Similarly, creole languages which make use of only a subset of the tense markings of their lexifier languages have been shown to have actually reanalyzed the earlier Tense systems as Aspectual systems. Quite obviously, neither of these or similar cases could reasonably be qualified as incomplete and not even as simplified systems, in spite of the fact that they exhibit a reduced array of surface devices, as compared to their diachronic ancestor languages.

The conclusion to be drawn from these considerations for the study of heritage languages is that in order to qualify the linguistic knowledge of HL learners as resulting from incomplete acquisition, it will not suffice to refer to a reduced set of surface expressions in comparison to those exhibited by

the corresponding full languages. Rather, this discussion should be based on analyses of the respective underlying grammatical systems and it should address the question of whether HL grammars are indeed incomplete or merely divergent from the ones they are compared to. In the latter case, it is furthermore crucial to determine whether the observed instances of simplification are specific to the language of HL speakers. If the same reductions are also attested in colloquial varieties of the full language, they are likely to indicate ongoing processes of change, possibly accelerated in the case of HL speakers because, as mentioned earlier, these are exposed primarily or exclusively to colloquial speech. Although this issue is not frequently addressed in studies of HL learners, some do provide evidence that certain HL properties occur in non-formal varieties of the full language as well. Polinsky (2006), for example, mentions the reduction of the case marking paradigm, the replacement of synthetic by analytic verb forms, and the loss of subject-verb agreement as phenomena which appear in the speech of Russian HL speakers but which are also attested in colloquial varieties of Full Russian. Finally, in characterizing the knowledge of HL speakers as incomplete, one also needs to determine whether the grammatical phenomena in which it differs from native grammars include aspects of narrow syntax. That this is indeed the case has been argued most recently by Cuza (2013) who analyzed Spanish word order, especially subject-verb inversion in embedded clauses, investigating US-born HL learners. He finds evidence for incomplete acquisition affecting “syntax proper”. This suggests that at least some HL learners are indeed L2 learners, in spite of the fact that they were exposed to this language from birth.

In sum, as far as acquisition processes are concerned, the most interesting problem raised by studies of HL speakers is whether some of these learners indeed acquire an only partial grammatical competence – provided incomplete acquisition is understood as resulting in a non-native kind of grammatical knowledge, possibly resembling that of second language learners. Note, however, that some of the individuals analyzed in HL studies were born in the countries of origin of their parents and were therefore exposed to the language of the host country only at later ages, sometimes as late as age 11. The group of HL learners thus includes simultaneous as well as early and late successive bilinguals; see Montrul (2011a). In order to be able to assess the roles which various factors play in instances of incomplete acquisition, comparing HL and L2 learners, these various types of acquisition should be investigated separately. The question thus is whether a HL indeed

“resembles in many aspects a second language acquired in adulthood as opposed to the first language acquired in childhood”, as stated by Rothman (2009: 157). Since HL speakers are typically exposed to the HL from birth but are claimed to receive considerably less input than learners of the full languages, the issue at stake is whether impoverished input can have similar effects on acquisition as delayed age of onset of acquisition. Montrul (2008: 7) rejects this possibility, stating that “even when an L1 may be partially lost or incompletely acquired in early bilingualism, it still retains the signatures of a first language because it was acquired during the Critical Period.” *Quod est demonstrandum!*

#### **4.2. Insufficient exposure to the target language as a possible cause of structural reduction**

There seems to be a broad consensus among researchers studying HLL that incomplete acquisition or structural simplification results from limited exposure to the target language and that this is the single most important property distinguishing HL learners from other learner types. Given that exposure to the L1 during the critical age ranges is a necessary but not sufficient condition for the development of a native competence, the potentially most important insight for a theory of language acquisition to be gained from studies of HL concerns the question of what constitutes a ‘sufficient condition’ for successful (bilingual) first language development with respect to exposure to the target language. Currently, however, we do not have an even approximate estimate of the threshold which needs to be attained in order for the amount of exposure to the target language to qualify as sufficient or adequate.

This is all the more surprising given that researchers investigating HL learners emphasize almost unanimously that the nature of the input is an important if not *the* crucial question to be answered by HL studies. Montrul (2011b: 187), for example, summarizes that

it is undeniable that successful language acquisition must start in childhood when the child is exposed to rich input (in quantity and quality) and has the opportunity to use the languages in a variety of meaningful situations. If one of these factors is less than optimal, linguistic deficits are very likely to appear throughout development and into adulthood. This is what happens to L2 learners and to heritage speakers.

This scenario refers to two of the three dimensions of the variation space constraining grammatical development, outlined at the end of section 3 above, and it emphasizes the claim according to which, for HL speakers acquiring a heritage language from birth or at least from very early on, it is the lack of “rich input” which is primarily responsible for deficiencies like the ones alluded to in the quote. But I have not been able to find, in the rapidly growing number of publications on HL learners, an insightful discussion of what might possibly constitute a sufficiently rich input allowing these learners to develop a native grammatical competence without such deficiencies. Yet it should be self-evident that only on the basis of detailed information about quantity and quality of input is it possible to assess the role of this factor for grammatical development and to determine whether or not impoverished input can have a similar effect as delayed onset of acquisition.

We are, of course, not totally ignorant with respect to these issues. But what we know we owe to research on children developing two first languages simultaneously, rather than to HL studies investigating bilinguals who are only partly successful in the acquisition of one of their languages. As mentioned in the summary of research on 2L1 development in section 3, one insight gained by these studies is that bilingual children are able to develop two native competences by mere exposure to the PLD even though the relative time of exposure to each of the two languages amounts at best to 50% of the total amount of child-directed speech. In fact, in most cases it will be less than that, given that the two languages are rarely distributed equally in the daily amount of child-directed utterances in a child’s linguistic environment. Another finding of 2L1 research is that bilinguals can attain native competence even in settings where one of their languages is weaker than the other one, i.e. if the two languages are not balanced with respect to their availability for the child; cf. Meisel (2007a) or Montrul (2008). In other words, if one of the factors mentioned in the quote from Montrul (2011b) is less than optimal, developmental deficiencies are not a necessary consequence. Rather, the LMC is a robust enough device to enable individuals to acquire native grammars in spite of reduced quantity of input. Yet there can hardly be any doubt that a minimum amount of exposure to the PLD is necessary for children to be able to develop a native competence, for, trivially, zero input will result in acquisition failure. The challenge thus is to determine in what way decreasing amounts of input affects language acquisition and whether it is possible to quantify a minimum threshold for successful acquisition of grammars.

One effect of limited exposure to the target language which is well documented concerns lexical development. Although it has been demonstrated that bilinguals' total vocabulary in both languages combined is virtually the same as in monolinguals, cf. Pearson, Fernández and Oller (1993), the vocabulary of bilinguals in each of their languages has been shown to be smaller than that of the respective monolinguals of the same age, and it also seems to develop more slowly, cf. Hoff, Core, Place, Rumiche, Señor and Parra (2012) who discuss such findings in some detail. However, the more pertinent question with respect to the issue of whether reduced input can cause incomplete acquisition is how limited exposure to PLD affects *grammatical* development.

Hoff *et al.* (2012) examined this question and concluded "that grammatical development in bilingual children is also a function of language exposure" (p.23). They compared Spanish–English bilingual to monolingual English children at ages 1;10, 2;1 and 2;6. To assess linguistic development, they applied the English MacArthur-Bates Communicative Development Inventory: Words and Sentences (CDI) and its Spanish counterpart, El Inventario del Desarrollo de Habilidades Comunicativas (IDHC). To measure grammatical complexity, these instruments rely on caregivers' indications of whether certain constructions are likely to have been produced by a child. Hoff *et al.* (2012) report that, across ages, the monolinguals in their study were more advanced in English than the bilinguals, and that the grammatical complexity of their English utterances increased faster, resulting in an increasing lag between the two learner types. Although the degree of developmental difference falls within the normal range of variation for monolingual children, what matters for our present concern is that the relative amount of language input in English correlated to the development of English. The English-dominant (70% or more of total exposure) bilinguals looked most like the monolinguals, followed by the balanced (50 – 60%) and the Spanish-dominant (10–30% exposure to English) bilinguals, in this order. As for the acquisition of Spanish, the Spanish-dominant bilinguals achieved better results than the balanced learners, and these did better than the English-dominant ones.

These findings clearly support the claim that reduced input affects the acquisition of grammar. Note, however, that the effects on grammatical development are smaller than on vocabulary development. More importantly, Hoff *et al.* (2012) did not find qualitative differences between monolinguals and bilinguals, i.e. bilinguals proceed through the same



sequence of developmental milestones as monolinguals. It remains to be seen whether these results can be corroborated by investigations analyzing children's speech rather than relying on the CDI or similar instruments. If this is the case, it is perhaps not implausible to assume that a significantly slower rate of grammatical development might entail reduced exposure to PLD during the relevant periods of heightened sensitivity. As a consequence, the available input might not suffice to trigger native-like development in certain domains of grammar. For the time being, however, this is merely a speculative remark.

Matters are, in fact, more complicated than what the preceding considerations suggest. They are based on the assumption that relative frequency of exposure to a language reflects more or less accurately the amount of input data available to the child. This, however, is far from obvious. Starting from the observation that the amount of speech which parents direct to children varies considerably across individuals and that bilingual children's amount of language input in each of their languages can therefore be equal to or larger than that to which monolingual children are exposed, De Houwer (2014) examined empirically the amount of input in Dutch received by monolingual and bilingual (Dutch-French) Belgian children at ages 1;1 and 1;8. She found no empirical support for the assumption that the amount of input is necessarily reduced for bilingual as compared to monolingual children. Her study is based on a comparison of various measures of maternal input frequency, i.e. frequency of utterances as well as of words, morphemes and syllables per utterance. The analysis revealed no group differences across these measures. De Houwer (2014) concluded that her study did not find evidence of reduced input in the bilingual as compared to the monolingual group. It did, however, confirm that there is extensive variation among mothers in speech rate per hour. In other words, children may hear a lower number of utterances addressed to them in the dominant language than in the one in which they receive proportionally less input.

As a preliminary summary we can retain that studies examining language acquisition in settings where children have only limited access to the PLD confirm the expected relation between amount of available input and grammatical development with respect to the rate of development. But even in cases where relative frequency of contact with a language does not attain more than 30%, this did not provide evidence for qualitative deficits in the attained grammatical knowledge, and certainly not for incomplete

acquisition. This is not really surprising, given that studies which analyzed ongoing acquisition processes of HL learners also concluded that, at least in some domains of grammar, these learners are able to develop a native competence in spite of less than optimal conditions for acquisition.

A recent example is the study by Flores and Barbosa (2014) who investigated the acquisition of the syntax of clitic placement in European Portuguese by 12 HL learners, age 7 to 15, living in Germany, comparing them to Portuguese monolinguals. The data analysis revealed that the young HL speakers used enclisis where proclisis is required by EP syntax; they eventually acquired proclitic placement, but later than the monolinguals. These authors argued that the HL learners take longer to acquire the grammar of clitic placement because they are exposed to reduced input, but that they proceed through the same developmental stages as the monolinguals.

These and similar findings confirm the claim according to which limited access to the target language can result in delays in the acquisition of grammars. But a threshold defining the minimum amount of input quantity for acquisition to be at all possible must also exist. In order to discover what this minimum might be, it is necessary to study cases where access to the target language is radically limited, i.e. where learners are exposed to much less PLD than those where the non-dominant language accounts for approximately 30% of the daily speech directed to children. Such a situation can arise with hearing children of deaf parents in their acquisition of spoken language. Flores and Barbosa (2014) report on studies (Schiff-Myers 1988) which suggest that these children require a minimum of 5-10 hours of weekly interaction with non-deaf people in order to develop a productive competence in their spoken L1. Note, however, that this does not reveal the total amount of input in communicatively meaningful contexts, nor does it allow us to conclude that these children did not acquire an incomplete grammatical competence.

Studies of HL learners should also be able to formulate testable hypotheses about the minimum input enabling children to acquire full rather than incomplete knowledge of the target grammar – provided they engage in empirical investigations of the course of acquisition by bilinguals in settings where access to one of the languages is drastically reduced. An example of this type of urgently needed research is the one by Silva-Corvalán (2003), analyzing the acquisition of the Spanish tense-modality-aspect (TMA) system by seven children. Importantly, she relied on longitudinal data from



two of these children, Nico and Bren. Nico was raised bilingually up to age 3; after that, the home language was English. From then on until age 5, his exposure to Spanish was then limited to approximately 3-4 hours per week. Bren's exposure to Spanish amounted to less than two hours per week during the entire period studied. By ages 5;1 – 5;6, they had both acquired a system of Spanish TMA morphology, but a simplified one, compared to that of monolinguals and to HL children who received more Spanish input. Interestingly, at ages 5;1-5;6, the two boys failed to produce some TMA forms which they used at ages 3;0 – 3;3. It thus seems that they not only acquired the Spanish TMA system incompletely but that their system was subject to attrition during the last two years.

These findings strongly suggest that incomplete acquisition of grammatical competence is likely to happen only when access to the PLD is dramatically reduced. As far as I can see, it is not possible, at this point, to give a more precise estimate of the quantity of input necessary for the development of a native or native-like grammatical competence. As we have seen, HL studies are, in principle, in a good position to make a significant contribution to the solution to this problem. However, in order to be able to do so, they need to gather detailed information about quantity and quality of input during sensitive phases for language development, and they must analyze the speech of learners and ideally also that of caretakers. To be clear, self-reports on the languages spoken with parents or siblings in childhood or reports by parents or teachers on the estimated relative frequency of exposure to the HL are not the kind of data required here. Nor do self-ratings of proficiency or language dominance provide the necessary data. In other words, studies on the acquisition of HL must meet the same methodological standards as first and second language acquisition research. In fact, in this particular case where the nature of input data is hypothesized to be a causal factor for deficient grammatical developments, it is furthermore necessary to gather information about the ambient languages, for it is of crucial importance to determine whether a feature allegedly indicating incomplete acquisition really emerged in the speech of HL learners or whether it was already present in the language used by their caretakers; cf. Domínguez (2009). Unless these requirements are met, claims to the effect that incomplete acquisition is caused by insufficient exposure to the target language are merely stipulations.

## 5. CONCLUSION

From the perspective of language acquisition research, HL learners are an extremely heterogeneous group, comprising a number of different learners types, each of these defined in terms of the major factors shaping course and ultimate attainment of acquisition. The question then is whether heritage languages constitute a genuine domain of acquisition research. This would be the case if at least some of those subsumed under the term 'heritage language learner' exhibited characteristics not shared by others. 'Incomplete' grammatical competence in one of the languages of bilinguals who acquired both languages from birth might qualify as such a learner-type specific property – provided incompleteness is understood as referring to non-native grammars. HL studies thus offer an almost unique opportunity to enhance our understanding of the human Language Making Capacity if they succeeded in revealing when this capacity fails in spite of the learner's exposure to the target language during sensitive phases for grammatical development.

The crucial question to be addressed here is what causes this kind of acquisition failure. Publications reporting on HL acquisition almost unanimously claim that incompleteness or simplification is due to insufficient exposure to the primary linguistic data of the target language. Unfortunately, we are largely ignorant as to what 'sufficient' or 'rich' as opposed to 'impoverished' or 'reduced' input might mean. In my view, specifying the quantitative threshold which needs to be attained in order to be able to develop native competence is the most important issue to which studies of HL learners can contribute. Note that this is would not only reveal important insights into the nature of the human Language Making Capacity, it is also of immediate practical relevance. If first generation immigrants want their children to acquire native competences in the language of the country of origin as well as in the language of the host society, and if they turn to linguists or psychologists for advice, it would hardly be satisfactory to suggest to them to provide "a sufficiently rich linguistic environment".



## REFERENCES

- Alarcón, Irma M. (2011): "Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behaviour within the noun phrase", *Bilingualism; Language and Cognition* 14, pp. 332-350.  
<http://dx.doi.org/10.1017/S1366728910000222>
- Albirini, Abdulkafi, Elabbas Benmamoun and Eman Saadah (2011): "Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production", *Studies in Second Language Acquisition* 33, pp. 273– 303.  
<http://dx.doi.org/10.1017/S0272263110000768>
- Andersen, Roger W. (ed.) (1983): *Pidginization and Creolization as Language Acquisition*. Rowley, MA: Newbury House.
- (1983): "Introduction: A language acquisition interpretation of pidginization and creolization", in Roger W. Andersen (ed.): *Pidginization and creolization as language acquisition*. Rowley, MA: Newbury House, pp. 1-56.
- Bylund, Emanuel (2009): "Maturational constraints and first language attrition", *Language Learning* 59, pp. 687-715.  
<http://dx.doi.org/10.1111/j.1467-9922.2009.00521.x>
- Chomsky, Noam (1981): "Principles and parameters in syntactic theory", in Norbert Hornstein and David Lightfoot (eds.): *Explanation in linguistics: The logical problem of language acquisition*. London: Longman, pp. 32-75.
- Cuza, Alejandro (2013): "Crosslinguistic influence at the syntax proper: Interrogative subject-verb inversion in heritage Spanish", *International Journal of Bilingualism* 17, pp. 71-96.  
<http://dx.doi.org/10.1177/1367006911432619>

- De Houwer, Annick (2014): "The absolute frequency of maternal input to bilingual and monolingual children: A first comparison", in Theres Grüter and Johanne Paradis (eds.), *Input and experience in bilingual development*. Amsterdam: John Benjamins pp. 37-58.
- Domínguez, Laura (2009): "Charting the route of bilingual development: Contributions from heritage speakers' early acquisition", *International Journal of Bilingualism* 13, pp. 271-287.  
<http://dx.doi.org/10.1177/1367006909339812>
- Flores, Cristina (2008): *A competência sintáctica de falantes bilingues luso-alemães regressados a Portugal: Um estudo sobre erosão linguística*. Universidade do Minho, Braga: Ph.D. dissertation.
- (2010): "The effects of age on language attrition: Evidence from bilingual returnees", *Bilingualism: Language and Cognition* 13, pp. 533-546.  
<http://dx.doi.org/10.1017/S136672890999054X>
- and Pilar Barbosa (2014): "When reduced input leads to delayed acquisition: A study on the acquisition of clitic placement by Portuguese heritage speakers", *International Journal of Bilingualism* 18, pp. 304-325.
- Foote, Rebecca (2010): "Age of acquisition and proficiency as factors in language production: Agreement in bilinguals", *Bilingualism: Language and Cognition* 13, pp. 99-118.  
<http://dx.doi.org/10.1017/S136672890999040X>
- Hoff, Erika, Cynthia Core, Silvia Place, Rosario Rumiche, Melissa Señor and Marisol Parra (2012): "Dual language exposure and early bilingual development", *Journal of Child Language* 39, pp. 1-27.  
<http://dx.doi.org/10.1017/S0305000910000759>
- Hyltenstam, Kenneth and Niclas Abrahamsson (2003): "Maturational constraints in second language acquisition", in Catherine J. Doughty and Michael H. Long (eds.): *Handbook of second language acquisition*. Oxford: Blackwell, pp. 539-588.  
<http://dx.doi.org/10.1002/9780470756492.ch17>
- McNeill, David (1966): "Developmental psycholinguistics", in Frank Smith and George A. Miller (eds.): *The genesis of language*. Cambridge, MA: MIT Press, pp. 15-84.


- Meisel, Jürgen M. (1977): “Linguistic simplification: A study of immigrant workers’ speech and foreigner talk”, in S. Pit Corder and Eddy Roulet (eds.): *The notions of simplification, interlanguages and pidgins and their relation to second language pedagogy*. Geneva: Droz, pp. 88-113.
- (2007a): “The weaker language in early child bilingualism: Acquiring a first language as a second language?”, *Applied Psycholinguistics* 28, pp. 495-514.
  - (2007b): “On autonomous syntactic development in multiple first language acquisition”, in Heather Caunt-Nulton, Samantha Kulatilake and I-hao Woo (eds.): *Proceedings of the 31<sup>st</sup> Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 26-45.
  - (2011): *First and Second Language Acquisition: Parallels and Differences*. Cambridge: Cambridge University Press.
  - (2013): “Sensitive phases in successive language acquisition: The Critical Period Hypothesis revisited”, in Cedric Boeckx and Kleantes Grohmann (eds.): *Handbook of biolinguistics*. Cambridge: Cambridge University Press, pp. 69-85.  
<http://dx.doi.org/10.1017/CBO9780511980435>
- Montrul, Silvina A. (2008): *Incomplete Acquisition in Bilingualism*. Amsterdam: John Benjamins.  
<http://dx.doi.org/10.1075/sibil.39>
- (2011a): “The linguistic competence of heritage speakers”, *Studies in Second Language Acquisition* 33, pp. 155– 161.
  - (2011b): “Morphological errors in Spanish second language learners and heritage speakers”, *Studies in Second Language Acquisition* 33, pp. 163– 192.  
<http://dx.doi.org/10.1017/S0272263110000720>
- Pearson, Barbara Zurer, Sylvia C. Fernández and D. Kimbrough Oller (1993): “Lexical development in bilingual infants and toddlers: Comparison to monolingual norms”, *Language Learning* 43, pp. 93–120.  
<http://dx.doi.org/10.1111/j.1467-1770.1993.tb00174.x>
- Pires, Acrisio and Jason Rothman (2009): “Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars”, *International Journal of Bilingualism* 13, pp. 211-238.  
<http://dx.doi.org/10.1177/1367006909339806>

- Polinsky, Maria (2006): "Incomplete acquisition: American Russian", *Journal of Slavic Linguistics* 14, pp. 191-262.
- (2011): "Reanalysis in adult heritage language: New evidence in support of attrition", *Studies in Second Language Acquisition* 33, pp. 305-328.  
<http://dx.doi.org/10.1017/S027226311000077X>
- Rothman, Jason (2009): "Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages", *International Journal of Bilingualism* 13, pp. 155-163.  
<http://dx.doi.org/10.1177/1367006909339814>
- Schachter, Jacquelyn (1989): "Testing a proposed universal", in Susan Gass and Jacquelyn Schachter (eds.): *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press, pp. 73-88.  
<http://dx.doi.org/10.1017/CBO9781139524544.007>
- (1990): "On the issue of completeness in second language acquisition", *Second Language Research* 6, pp. 93-124.
- (1996): "Maturation and the issue of Universal Grammar in second language acquisition", in William C. Ritchie and Tej K. Bhatia (eds.): *Handbook of second language acquisition*. San Diego: Academic Press, pp. 159-193.
- Schiff-Myers, Naomi B. (1988): "Hearing children of deaf parents", in Dorothy Bishop and Kay Mogford (eds.): *Language development in exceptional circumstances*. Edinburgh: Churchill Livingstone, pp. 47-61.
- Schlyter, Suzanne (1993): "The weaker language in bilingual Swedish-French children", in Kenneth Hyltenstam and Åke Viberg (eds.): *Progression and regression in language: Sociocultural, neuropsychological and linguistic perspectives*. Cambridge: Cambridge University Press, pp. 289-308.
- Sekerina, Irina A. and John C. Trueswell (2011): "Processing of contrastiveness by heritage Russian bilinguals", *Bilingualism; Language and Cognition* 14, pp. 280-300.  
<http://dx.doi.org/10.1017/S1366728910000337>
- Seliger, Herbert W. (1978): "Implications of a multiple critical periods hypothesis", in William C. Ritchie (ed.): *Second language acquisition research*. New York: Academic Press, pp.11-19.

- Silva-Corvalán, Carmen (2003): “Linguistic consequences of reduced input in bilingual first language acquisition”, in Silvina Montrul and Francisco Ordóñez (eds.): *Linguistic theory and language development in Hispanic languages*. Somerville, MA: Cascadilla Press, pp. 375-397.
- Slobin, Dan I. (1985): “Crosslinguistic evidence for the Language-Making Capacity”, in Dan I. Slobin (ed.): *The crosslinguistic study of language acquisition. Vol 2: Theoretical issues*. Hillsdale, NJ: Lawrence Erlbaum, pp. 1157-1256.
- Sorace, Antonella (1993): “Incomplete vs. divergent representations of unaccusativity in non-native grammars of Italian”, *Second Language Research* 9, pp. 22-47.  
<http://dx.doi.org/10.1177/026765839300900102>
- (2000): “Syntactic optionality in L2 acquisition”, *Second Language Research* 16, pp. 93-102.
  - (2005): “Selective optionality in language development”, in Leonie Cornips and Karen P. Corrigan (eds.): *Syntax and variation: Reconciling the biological and the social*. Amsterdam, Philadelphia: John Benjamins, pp. 55-80.  
<http://dx.doi.org/10.1075/cilt.265>







# The English Phonological Skills of Latino Spanish-English Dual-Language Preschoolers Living in Los Angeles: Implications for Practice

Simona Montanari  
Kaveri Subrahmanyam  
Marlene Zepeda  
Carmen Rodríguez  
*California State University, Los Angeles*

**Abstract:** This study examines the English phonological skills of ten Spanish L1/English L2 bilingual preschoolers from Los Angeles, focusing on (1) phonetic inventory complexity, (2) overall consonant accuracy, (3) accuracy of production of early-, middle-, and late-developing sounds, and (d) interaction between phonological systems. The results indicate that English inventories were commensurate to those of same-age monolinguals, yet accuracy levels were lower. Spanish sounds appeared in English inventories, and English-only sounds were produced with lower accuracy than sounds present in both languages. These results validate previous findings and suggest that phonological systems interact in development, bootstrapping some phonological skills while at the same time delaying others. We discuss the implications of our findings for speech pathologists and preschool educators.

**Keywords:** Phonological development, phonological skills, Spanish-English bilingualism, preschoolers.



## 1. INTRODUCTION

The United States has gone through a major demographic shift in the last twenty years. According to the 2010 U.S. Census, Latinos are now the second largest ethnic group in the US ([www.census.gov](http://www.census.gov)). In California, Spanish-speaking Latinos represent close to one third of the entire population and 40% of Los Angeles County residents speak Spanish. Latino children make up the largest group of children in the US who speak a language other than English in the home. Many of these children begin learning English when entering the school system, while they are still developing Spanish, and are called *Dual Language Learners* (DLLs) (Paradis, Genesee, and Crago, 2011)<sup>1</sup>.

Despite the increasingly large number of Spanish-speaking DLLs in the US, little is known about typical L1 and L2 phonological skills during the preschool years. Because of this limited knowledge, assessment of bilingual children is often based on norms obtained from monolingual speakers (Gutiérrez-Clellen, 1996) and at times leads to bilingual children being over-diagnosed or under-diagnosed with speech sound disorders. Over-diagnosis can result in unnecessary speech therapy and might have educational consequences for children who are removed from the classroom to receive these services (Lidz and Peña, 1996), whereas under-diagnosis can delay crucial intervention, resulting in prolonged speech problems and missed educational opportunities. Additionally, as the proportion of DLLs increases in the US, it is likely that the number of children referred to or diagnosed with speech and/or language disorders will also increase substantially over the next decades. Clearly, there is an urgent need to understand more about typical bilingual phonological development and, therefore, have a valid basis for assessment and intervention in this growing population.

Understanding the phonological skills of DLLs has important educational implications for their school success and achievement. Phonological development is a strong predictor of both lexical development (Stoel-Gammon, 2011) and future literacy abilities (decoding, reading comprehension, and spelling) in both monolingual and bilingual students (National Center for Family Literacy, 2008). Currently, Spanish-speaking DLLs lag behind English-speaking students on measures of school readiness

---

<sup>1</sup> The terms ‘Dual Language Learners’ and ‘bilingual children’ will be used interchangeably throughout this article.

when starting kindergarten (National Task Force on Early Childhood Education for Hispanics, 2007), and, overall, Latinos perform far below their white counterparts on statewide assessments in reading and language arts (National Assessment of Educational Progress, 2011). Given the link between phonological and literacy development, information on the phonological skills of young DLLs can provide important insights on the best approaches to teaching young DLLs.

### **1.1. Phonological development in Spanish-English bilingual preschoolers**

Recent research has begun to examine phonological development in Spanish-English bilingual preschoolers (Fabiano-Smith and Barlow, 2010; Fabiano-Smith and Goldstein, 2010a, 2010b; Gildersleeve-Neumann, Kester, Davis, and Peña, 2008). For example, Fabiano-Smith and colleagues have examined phonetic inventories (i.e., number of sounds they can produce) (Fabiano-Smith and Barlow, 2010), overall sound accuracy (Fabiano-Smith and Goldstein, 2010a, 2010b), accuracy of production of early-, middle- and late-developing sounds (Fabiano-Smith and Goldstein, 2010a), and signs of interaction between Spanish and English phonological systems (Fabiano-Smith and Goldstein, 2010b) in 24 three-to-four-year-old Spanish-English bilingual and monolingual preschoolers (eight monolingual English and eight Spanish-English bilingual children from Philadelphia, Pennsylvania, and eight monolingual Spanish-speaking children from Querétaro, Mexico). Overall, these studies' findings suggest that bilingual preschoolers might demonstrate differing phonological skills depending on the phonological construct being investigated. In particular, their phonetic inventories in each language might be comparable to those of monolinguals (Fabiano-Smith and Barlow, 2010), while, at the same time, their accuracy of vowel and consonant production may be lower than in monolinguals (Fabiano-Smith and Goldstein, 2010a, 2010b). As Fabiano-Smith and Goldstein (2010b: 25-26, emphasis in original) note, "lower accuracy in bilingual production does not indicate delayed or disordered acquisition in a clinical sense"; on the contrary, it shows "a slower, *but nonetheless relatively similar*, rate of acquisition [than in monolinguals]" at some points in development. Similar results were reported by Gildersleeve-Neumann et al. (2008), who examined the English phonological skills of 33 monolingual and Spanish-English bilingual three-to-four-year-olds over a school year (10 English-only, 20 predominantly-English, and 3 balanced bilingual children, all

residing in Texas). The authors found a slightly larger English error rate for bilingual children when compared to English monolinguals. However, the rate of change was similar across all children studied and with exposure to and practice in English, bilingual children achieved similar outcomes to monolinguals.

While these studies are important, they are only a first step in bettering our understanding of typical Spanish-English bilingual phonological development. No study, to date, has examined the phonological skills of bilingual children living in California, the state that educates one third of all DLLs in the U.S. (approximately 5.3 million children or 9 to 10% of the entire U.S. school population, National Clearinghouse for English Language Acquisition, 2011). Spanish-speaking children in California are not only increasing but they are soon predicted to become the majority language group in California. According to the California Department of Education ([www.cde.ca.gov/ds/](http://www.cde.ca.gov/ds/)), for instance, Spanish-speaking children made up 31% and 37% of incoming kindergarteners in California and Los Angeles, respectively, in the year 2012-2013. In addition, it has been estimated that the percentage of Spanish-speaking children is double this figure for the zero-to-five-year-old population (Children Now, 2014). Given these numbers, it is important that Spanish-speaking children from California be included in any attempt to examine Spanish-English bilingual phonological development.

Previous studies have focused primarily on bilingual speakers of Puerto Rican, Dominican, and Mexican Spanish living in Pennsylvania and Texas. Bilinguals growing up in regions like Pennsylvania, which are geographically distant from Spanish-speaking Mexico and Central America, might live in bilingual environments that are more subtractive and actually better supportive of English than California. Also, the subjects of previous studies spoke English and Spanish varieties that are different from those spoken in California. Although Latinos are often viewed as a homogenous group of Spanish speakers, Latino subgroups and even Mexican-descent individuals in different parts of the U.S. (e.g. Texas vs. California) differ substantially when it comes to dialectal features (Whitley, 2002) and developmental patterns attested in one group or dialect might not be necessarily found in another (Macken and Barton, 1980, and Eilers, Oller, and Benito-García, 1984). Some of the most striking phonological differences between Mexican and Caribbean Spanish (in particular, Puerto Rican and Dominican Spanish), for example, include (a) the strong, sibilant

pronunciation of syllable- and word-final /s/ in Mexican Spanish ([espera] for *espera*) as opposed to its aspiration in the Puerto Rican and Dominican varieties ([ehpera]); (b) the production of /r/ as an alveolar trill in Mexican Spanish ([pero] for *perro*) as opposed to its velarization [x] or uvularization [χ] in Puerto Rican Spanish ([pexo] or [peχo]); and (c) the production of final /n/ as alveolar in Mexican Spanish ([pan] for *pan*) as opposed to velar in Caribbean Spanish ([paŋ]) (Whitley, 2002).

Goldstein and Gildersleeve-Neumann (2012) discuss the difficulties of studying Spanish-English bilingual phonological development in populations that speak different Spanish and English varieties. The authors argue that, in this case, the phonology of one language not only may influence phonological production in the other language, but dialect-specific features of one language may transfer to the other language. For example, Spanish speakers who exhibit final /s/ deletion in Spanish as a typical dialectal feature (speakers of “radical dialects” [Guitart, 1996] as Puerto Rican and Dominican Spanish) might also show this feature in English, realizing words such as /mæs/, *mass*, without final /s/ as in [mæ]. On the other hand, speakers whose Spanish variety does not have this characteristic (speakers of “conservative dialects” [Guitart, 1996] as Mexican Spanish) might not display this same error. Therefore, since the particular dialectal features of the L1 variety might affect L2 phonological development, it is important to examine whether English phonological development differs among speakers of different Spanish varieties (e.g. Puerto Rican and Mexican Spanish). Such a study is necessary to validate the findings of extant work and will inform us more broadly about the course of L2 phonological development among Spanish-English DLLs.

Also relevant here is the proposal that “atypical” or “monolingual-deviant” phonological patterns among DLLs might result from the *interaction* between the bilinguals’ language systems (Paradis and Genesee, 1996). Researchers focusing on *early bilinguals* have shown indeed that although these children develop separate language – and phonological – systems (see Genesee and Nicoladis, 2007, for a review), these systems do not develop autonomously but interdependently. This interdependence might lead early bilinguals to *transfer* elements from a more dominant to a weaker language (Paradis and Genesee, 1996), facilitating the acquisition of some features in this weaker language (*acceleration hypothesis*, Paradis and Genesee, 1996), for example sounds that are phonetically similar cross-linguistically (*positive transfer*, Goldstein and Bunta, 2012). However,

interaction might also result in *negative transfer* (Goldstein and Bunta, 2012) when sounds in one language do not have phonetically similar sounds in the other, therefore slowing down phonological development in this latter language (*deceleration hypothesis*, Paradis and Genesee, 1996). While researchers agree that such cross-language interaction is a typical feature of early bilingual development, past studies of Spanish-English bilinguals have found evidence of both acceleration (Goldstein and Bunta, 2012) and deceleration (Fabiano-Smith and Goldstein, 2010a, 2010b; Gildersleeve-Neumann et al., 2008).

At the same time, an opposite claim has been made for speech sound acquisition among *older L2 learners*. In particular, Flege's (1995) Speech Learning Model (SLM) predicts that L2 sounds that do not have equivalents in L1 are easier to acquire than sounds that are similar between L1 and L2. This is because an L2 sound without an L1 equivalent will be more easily perceived as "different" by the learner and will more likely result in the formation of a new sound category. Although, according to the SLM, new L2 sound categories are more easily established in childhood due to the less robust nature of L1 categories, Flege's proposal directly contradicts Paradis and Genesee's (1996) and Goldstein and Bunta's (2012), positing that sound similarity *interferes* rather than helps L2 phonological development. Therefore, given these opposing proposals and the fact that our participants are both early bilinguals and L2 learners, more studies are needed to examine the extent and consequences of interaction between phonological systems and whether sound similarity prevents or rather aids the development of L2 phonology.

### 1.2. The present study

The present study examined the English phonological skills of a group of Mexican-origin Spanish L1/English L2 preschoolers from Los Angeles, California, using the same measures employed in previous studies (Fabiano-Smith and Barlow, 2010; Fabiano-Smith and Goldstein, 2010a, 2010b). However, unlike those studies, which assessed phonological skills on the basis of a limited number of words, we used an extensive list of 119 lexical items (containing at least three instances of all English consonants word-initially and word-finally) to ensure that the participants had ample opportunity to produce all English consonants in different word positions. The study then examined the following research questions:



1. How complex are the bilingual children's English inventories? Are they commensurate with monolingual inventories or sparser due to bilingual learning?
2. Are accuracy levels different for early-, middle-, and late-developing sounds? Are the bilingual children's accuracy levels similar or lower than in same-age English monolinguals and similar-background bilingual children?
3. Are there signs of interaction between the children's Spanish and English phonological systems? In particular, do English inventories display transfer from Spanish, e.g., Spanish-specific sounds? Are accuracy rates lower for sounds not shared by Spanish and English, that is, sounds that only exist in the English inventory like /v/, /z/ or /θ/)?

## 2. METHODOLOGY

### 2.1. Participants

The participants included ten typically developing Spanish L1/English L2 bilingual children, with a mean age of 3;10 (age range: 3;4-4;4); there were five boys and five girls. All children came from low-income, Mexican-origin families who lived in Los Angeles County and were enrolled in a Head Start program during the time of the study. In order to determine the children's length and amount of exposure, use, proficiency, and preference in each language, a detailed parent questionnaire was administered in Spanish. As shown in Table 1, all children were reported to have Spanish as their L1, to use Spanish most of the time when interacting with parents and siblings and to have Spanish as their dominant and preferred language. At the same time, all children were early bilinguals as they had begun learning English around age three or earlier through older siblings, other family members, and the media. At the time of the study, all children had been formally and consistently exposed to English for seven months in the Head Start program, which they attended five days a week for four hours a day. Although English was the language that was used and encouraged in the classroom, the majority of teachers and classroom assistants were Spanish/English bilinguals from the community, and they did address children in Spanish at times.

LANGUAGE EXPOSURE (3= Began at birth; 2= Started between 1 and 2 years; 1= Started at age 3)		
Exposure to Spanish (SD) <b>3.0 (0)</b>	Exposure to English (SD) <b>1.3 (0.4)</b>	
LANGUAGE USE (4= Spanish only; 3= Mostly Spanish; 2= Mostly English; 1= English only)		
Language spoken to child by parents (SD) <b>3.9 (0.3)</b>	Language spoken by child to parents <b>3.6 (0.5)</b>	Language spoken with siblings (SD) <b>3.4 (0.5)</b>
PROFICIENCY AND PREFERENCE (3= Spanish; 2= Both, 1= English)		
Child's stronger language (SD) <b>2.9 (0.3)</b>	Child's preferred language (SD) <b>2.6 (0.5)</b>	

**Table 1.** Participants' language exposure, use, proficiency, and preference as reported by parents.

In order to make sure that the children had English proficiency levels that were commensurate to those of the participants of previous studies, the children's MLUm (i.e. mean length of utterance in morphemes) was calculated on the basis of conversational speech samples collected at the Head Start site by a bilingual researcher (who only addressed the children in English). The children's mean MLUm was 2.88 ( $SD = 0.86$ ). This value suggests that the English produced by the children was in the typical range for their age and bilingual status (when compared to the values given by the Bilingual S/E SALT program, Miller and Iglesias, 2006, for English productions of bilingual children). This value also indicates that the participants had English proficiency levels comparable to those of the children in Fabiano-Smith and colleagues' studies, who reported a mean MLUm value of 2.70 for their 8 bilingual subjects ( $SD = 0.91$ ).

## 2.2. Data collection

Single word samples were collected from each child by the same research assistant who had collected the conversational speech samples. Word samples were elicited using a list of 119 stimulus words (available from the authors). The stimulus words were adapted from Barlow's (2003b) *Assessment of English Phonology* and contained at least three instances of all English consonants both in word-initial and final position. Pictures were created for each target word. Each item was then elicited by asking the

child to name the object depicted in the picture. In the case of no response, children were given prompts or were allowed to repeat the answer provided by the examiner (as in previous studies).

Each sample was recorded using an Edirol R-09HR High-Resolution WAVE/MP3 recorder and phonetically transcribed in IPA narrow transcription by a graduate student and the first author, who are both phonetically trained and bilingual in Spanish and English. Disagreements on sounds were discussed by listening to the recordings several more times. If agreement was not reached, the affected sounds were excluded from the analyses. This resulted in 97.6% of identically transcribed data that was then subject to phonological analyses.

### 2.3. Analyses

#### 2.3.1. Phonetic inventories

In order to assess which sounds each child could produce, phonetic inventories were derived from the children's productions, following the same procedure traditionally employed to study this issue (Dinnsen, Chin, Elbert, and Powell, 1990; Fabiano-Smith and Barlow, 2010; Stoel-Gammon, 1987). A sound was included in a child's inventory if it was produced *accurately* in at least two different words. That is, if the child produced /p/ accurately in *pen*, [pɛn], and *pot*, [pɒt], then /p/ was entered in the child's inventory (even if errors existed with this sound in other words). Then, in order to measure the inventories' complexity, the children's phonetic inventories were assigned to one of the five complexity levels of Dinnsen et al.'s (1990) implicational hierarchy of phonetic distinctions for English (shown in Table 2 together with the results) and compared to the levels attained by same-age English monolinguals and similar-background bilingual children. This typology classifies children's phonetic skills on the basis of sound *classes* rather than individual sounds present in inventories. Generally, less complex inventories will consist of earlier-developing sounds and will display fewer sound classes. More complex inventories will have more distinctive features and will consist of sounds that are usually acquired later.

In order to further assess whether DLLs acquire the phonology of their L2 in a simple-to-complex fashion, inventories were also examined for the presence of early-, middle-, and late-developing sounds (EML) (Shriberg, 1993). Shriberg (1993) proposed and validated this construct by comparing his data on consonant mastery in 64 English monolingual children with

phonological disorders with studies of typically-developing English-speaking children. He found that both children with phonological disorders and typically developing children displayed a similar order of acquisition for the 24 English consonants. He therefore divided these sounds into three groups: early- (/m, b, j, n, w, d, p, h/), middle- (/t, ɲ, k, g, f, v, tʃ, dʒ/) and late-developing sounds (/ʃ, ð, s, z, θ, l, ɹ/)<sup>2</sup> according to their accuracy of production. As argued by Fabiano-Smith and Goldstein (2010a), the EML construct might be a more valid measure of bilingual phonological skills than overall accuracy because it makes a direct link between developmental order and consonant mastery.

### 2.3.2. Overall consonant accuracy and accuracy of EML sounds

As in Fabiano-Smith and Goldstein (2010a, 2010b), we did not limit our analyses to phonetic inventories but we also examined overall consonant accuracy and accuracy of EML sounds. Overall consonant accuracy was calculated in terms of Percent of Consonants Correct (PCC) (Shriberg, Austin, Lewis, McSweeney, and Wilson, 1997). PCC indicates the percentage of consonants that the child has produced correctly out of the total number of consonants targeted. Therefore, contrary to the phonetic inventory construct, which measured which sounds the children could produce (minimally in two different words), PCC assessed the *overall accuracy of consonant production*. In order to examine whether accuracy levels were different for EML sounds, PCC was also calculated for these different categories and the results were directly compared to those obtained in previous studies.

### 2.3.3. Interaction between phonological systems

In order to assess whether there was interaction between the children's Spanish and English phonological systems, we examined (a) whether English words—and thus inventories—displayed Spanish-specific sounds (for example, whether English item *hug*, /hʌg/, was produced with the Spanish-specific sound /x/ in place of /h/ as in [xʌg]); and (b) whether accuracy rates were different for (1) sounds shared by both Spanish and English, that is, sounds that are “phonetically similar” (Flege, 1981) and present in both languages (/p/, /b/, /t/, /d/, /k/, /g/, /m/, /n/, /f/, /s/, /ð/ or [ð]<sup>3</sup>, /tʃ/,

<sup>2</sup> The sound /ʒ/ was excluded from analysis by Shriberg (1993) and by Fabiano-Smith and Goldstein (2010a), therefore it was also excluded from the accuracy analyses in the present study.

<sup>3</sup> /ð/ is a phoneme in English but an allophone [ð] of /d/ in Spanish.

/l/, /w/, /j/) and for (2) unshared sounds, that is, sounds that exist only in English (/ŋ/, /v/, /ʒ/, /z/, /ʃ/, /θ/, /h/, /dʒ/, /ɹ/) or in Spanish (/ɲ/, [β], /x/, [y], /r/, /r/). Notice that although the voiced interdental fricative is traditionally considered a phoneme in English but an allophone of /d/ in Spanish, /ð/ and [ð] were coded as ‘similar’ in this study following Fabiano-Smith and Goldstein’s (2010a) methodology. The authors argue indeed that the Spanish spirant can be considered as the underlying form of /d/ given that it occurs in more phonetic contexts than the voiced stop. This frequency of occurrence provides the child with a greater number of opportunities to produce the spirant, making this sound more fundamental to the Spanish phonemic inventory than the voiced stop (Barlow, 2003a). Mean accuracy and standard deviations were calculated for shared and unshared sounds and the results were compared to those obtained by Fabiano-Smith and Goldstein (2010b) for monolingual and bilingual peers.

### 3. RESULTS

#### 3.1. Phonetic inventories

The subjects’ English phonetic inventories are shown in Table 2. As can be seen, nine children had the most complex inventory type (Level E), while only one child had a Level D inventory. The early-developing sounds /m, b, j, n, w, d, p, h/ were produced by all children. Among middle-developing sounds, only five out of eight consonants were in all inventories. These included the stops /t, k, g/ and the affricates /tʃ, dʒ/. Late-developing sounds were the most challenging, and, in a pattern similar to monolingual English speakers, no child produced all the consonants in this category. Although all children exhibited complex English phonetic inventories, there was substantial variation across children. One child, CH06, had all English consonants except for the post-alveolar voiced fricative /ʒ/ (as in *treasure*, [tɹɛʒə]) (for a total of 23 sounds). On the other hand, subject CH03, who displayed a Level D inventory, had only 18 sounds, lacking a stridency distinction (that is, she did not have both /s/ and /θ/ or /ð/ and /z/ in her inventory). Children also differed in the number of Spanish sounds that appeared in their English productions. This issue is discussed in the section “Cross-language interaction” below.

	Early-developing Sounds	Middle-developing Sounds	Late-developing Sounds	Spanish Sounds	
CH01	m b j n w d p h	t k g f tʃ dʒ	ʃ ø s l ɹ	β ʎ <sup>1</sup>	LEVEL E
CH02	m b j n w d p h	t ŋ k g tʃ dʒ	ø s z θ l ɹ		LEVEL E
CH03	m b j n w d p h	t ŋ k g f tʃ dʒ	ʃ s l	β ʎ <sup>1</sup> r x	LEVEL D
CH04	m b j n w d p h	t k g f v tʃ dʒ	ʃ ø s z l ɹ		LEVEL E
CH05	m b j n w d p h	t ŋ k g f tʃ dʒ	ʃ ø s z θ l ɹ	β r	LEVEL E
CH06	m b j n w d p h	t ŋ k g f v tʃ dʒ	ʃ ø s z θ l ɹ		LEVEL E
CH07	m b j n w d p h	t ŋ k g f tʃ dʒ	ʃ ø s z θ l ɹ	r	LEVEL E
CH08	m b j n w d p h	t ŋ k g f tʃ dʒ	ʃ s l ɹ		LEVEL E
CH09	m b j n w d p h	t ŋ k g f v tʃ dʒ	ʃ ø s θ l ɹ		LEVEL E
CH10	m b j n w d p h	t ŋ k g f tʃ dʒ	ø s z θ l ɹ		LEVEL E

LEVEL A Nasals (m, n), stops (b, d, t), and glides (w, j) with no voice/manner contrast

LEVEL B Voice distinction added (p, b, t, d, k, g)

LEVEL C Manner of articulation distinction added, i.e. fricatives (f, v, s) and/or affricates (tʃ, dʒ)

LEVEL D One liquid consonant (l, ɹ) added

LEVEL E Laterality (l, ɹ) and/or stridency (θ, s or ø, z) distinction added

<sup>1</sup>This is the devoiced assibilated allophone of the Spanish flap, a common consonantal process in Mexican Spanish (Whitley, 2002).

Table 2. English inventories and their complexity levels according to Dinnsen et al.'s (1990) implicational hierarchy of phonetic distinctions for English.

### 3.2. Overall consonant accuracy and accuracy of Early-, Middle-, and Late-developing sounds

Table 3 shows the means and standard deviations for overall accuracy and accuracy of EML sounds for our subjects and the bilingual and monolingual children in Fabiano-Smith and Goldstein (2010a). In our study, mean PCC varied significantly from category to category but it was never below 60%. A Friedman test indicated that differences in accuracy between EML sounds were statistically significant ( $\chi^2(2, N = 10) = 15.800, p = .001$ ). In order to determine which sound classes differed from one another on accuracy, the Wilcoxon tests were performed. The results of the Wilcoxon tests indicated a significant difference in accuracy between early- and middle-developing sounds ( $z = -2.78, p = .0054$ ), middle- and late-developing sounds ( $z = -2.191,$

$p = .028$ ), and early- and late-developing sounds ( $z = -2.803$ ,  $p = .005$ ). In other words, accuracy within the three categories largely decreased from early- to late-developing sounds. Conversely, variation across children for the three sound classes consistently increased from early- to late-developing sounds. This means that accuracy was more uniform for easy-to-articulate sounds like stops and nasals (/b, d, p, m, n/) while children differed in how good they were at reproducing more complex sounds like liquids (/l, ɹ/) and fricatives (/v, ʃ, ð, s, z, θ/) which are produced by forcing air into two semi-occluded articulators and are typically acquired late.

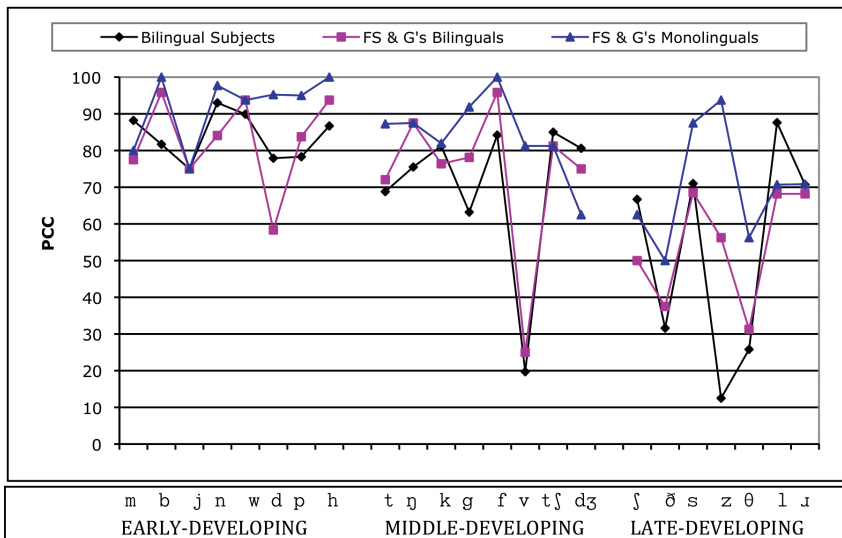
	Overall PCC		PCC Early		PCC Middle		PCC Late	
	M	SD	M	SD	M	SD	M	SD
Bilingual Subjects	<b>72.64</b>	<b>11.18</b>	<b>84.56</b>	<b>6.29</b>	<b>70.97</b>	<b>7.00</b>	<b>62.39</b>	<b>14.28</b>
FS and G's Bilinguals	72.31	12.45	82.76	8.65	74.08	13.24	63.91	20.88
FS and G's Monolinguals	84.10	8.20	93.93	3.97	86.53	7.60	74.10	16.07

**Table 3.** Means and standard deviations for overall percentage of consonants correct (PCC) and PCC for English early-, middle-, and late-developing sounds for the bilingual subjects and the bilingual and monolingual children in Fabiano-Smith and Goldstein (2010a).

When compared to the results of Fabiano-Smith and Goldstein (2010a), whose subjects lived in Pennsylvania, our findings are strikingly similar. In both studies, overall PCC was approximately 72% for the bilinguals, while PCC for EML sounds was approximately 83%, 72%, and 63% respectively. In both studies, the bilinguals had lower accuracy rates than the monolinguals, although, in Fabiano-Smith and Goldstein’s study, this difference was statistically significant only for early-developing sounds. Gildersleeve-Neumann et al. (2008) also found that their bilingual subjects in Texas had lower accuracy rates for consonants than the monolinguals; in addition, PCC for the English-dominant bilingual children was also around 70% as in our data, while PCC for monolinguals was around 80% as in Fabiano-Smith and Goldstein’s study. While we cannot assess whether similarities and differences between our findings and those of previous studies are statistically significant, a simple comparison of accuracy levels is useful not only because it validates previous findings but because it can ultimately help determine typical accuracy levels for EML sounds among Spanish L1/English L2 preschoolers.



Accuracy for individual sounds is reported in Figure 1. For comparison reasons, the results obtained by Fabiano-Smith and Goldstein (2010a) for their monolingual and bilingual subjects are also plotted in the same figure. As can be seen, our findings mirrored Fabiano-Smith and Goldstein’s for the bilinguals for the majority of segments, including /j, w, p, h/ among early-developing sounds, /t, k, v, tʃ/ among middle-developing ones, and /ð, s, θ, ɹ/ among late-developing sounds. In both studies, /d/ (early), /v/ (middle), /θ/ and /ð/ (late) were the most difficult consonants for the bilingual preschoolers (i.e. words containing those segments – *dog, vase, think, and there* – displayed errors), and the discrepancy between monolinguals and bilinguals was particularly evident for the latter three sounds. In addition, the children in this study produced errors with the alveolar voiced fricative /z/ as well (*zoo, zipper* and *zebra* were often produced as [su], [sɪpə] and [siba]). Taken together, these findings suggest that, for some sounds, Spanish L1/English L2 preschoolers might exhibit phonological skills that are commensurate to those of monolingual peers; for other, specific consonants, however, they might be demonstrating slower acquisition. The possible reasons behind this “un-uniform” development are discussed in the next section, which focuses on the interaction between children’s phonological systems.



**Figure 1.** Percentage of consonants correct (PCC) for English early-, middle-, and late-developing sounds for the bilingual subjects and the bilingual and monolingual children in Fabiano-Smith and Goldstein (2010a).



### 3.3. Cross-language interaction

#### 3.3.1. Spanish sounds in English productions

Table 2 shows that 40% of the children displayed transfer of Spanish-specific sounds to their English productions. Three children consistently produced the Spanish bilabial voiced fricative [β] in place of /v/, a sound that does not exist in Spanish (words like *van*, /væn/, or *vase*, /veis/, were produced as [βæn] and [βeis]). Although the bilabial fricative does not occur word-initially after a pause in Standard Spanish, word-initial spirantization is common in Mexican Spanish, even in compulsory stop environments (Amastae, 1995; Lozano, 1979; Macken and Barton, 1980). Macken and Barton (1980), for instance, found that 30%-40% of word-initial /b, d, g/ were spirantized by Mexican Spanish-speaking adults; similarly, both monolingual Spanish-learning children (Macken and Barton, 1980), and multilingual children learning Spanish in Los Angeles (Mayr and Montanari, in press), predominantly produced the bilabial fricative in compulsory stop contexts. Therefore, it is not surprising that the children in this study also made use of this process.

Three participants also replaced the English approximant /ɹ/ with the Spanish flap /r/ (e.g. words like *rain*, /ɹeɪn/, *present*, /pɹeɪzənt/, and *three*, /θɹi/, were produced as [reɪn], [pɹeɪzənt] and [θri] respectively). Two more children substituted the approximant word-finally with the devoiced assibilated version of the Spanish flap [ɾ̥] (e.g. *there*, /ðeɪ/, became [dɛɾ̥]). This is a common consonantal process in Mexican Spanish and other Latin American Spanish varieties (Perissinotto, 1975, reports that 73.5% of Mexico City youth uses R-assibilation), whereby flaps and trills are produced with frication (i.e. assibilated) and often devoiced in word-medial and word-final position (Whitley, 2002). Furthermore, one child (CH03) —who exhibited the least complex inventory and the largest amount of transfer from Spanish to English— produced the voiceless velar fricative /x/ in place of /h/ in at least two instances (e.g. *hug*, /hʌg/, became [xʌg]). Finally, we also found some instances of de-aspirated stops when they should have been aspirated (e.g. [pɪg] in place of [pʰɪg], *pig*). In sum, transfer from Spanish was evident, although children varied in the extent they used Spanish sounds in their English productions and/or exhibited systematic modification of low-level phonetic characteristics like aspiration.

3.3.2. Accuracy for shared and unshared sounds

Accuracy for sounds shared by both Spanish and English — “phonetically similar” (Flege, 1981)— and for unshared sounds is shown in Table 4. Results from Fabiano-Smith and Goldstein’s study (2010b) are also reported, both for the bilingual and the monolingual preschoolers. A Wilcoxon signed ranks test indicated that there was a significant difference between PCC of shared and unshared sounds among our bilingual subjects ( $z = -2.846, p = .004$ ). This means that, similarly to the bilingual children in Fabiano-Smith and Goldstein’s study, the participants of this study demonstrated significantly higher accuracy on sounds that were common to both Spanish and English like stops, nasals, glides, some fricatives (/f, s/), the affricate /tʃ/, and the liquid /l/, than on sounds that were present in only one language. In addition, both our subjects and the bilinguals in Fabiano-Smith and Goldstein’s study displayed lower accuracy rates than the monolinguals for both types of sounds. Notice that our subjects were not less accurate on unshared sounds simply because they tended to be late-developing. Among late-developing, shared sounds, accuracy was indeed 63.4%. Among late-developing, unshared sounds, on the other hand, accuracy was 41.1% (the results of a Wilcoxon test indicates that this difference was significant:  $z = -1.86, p = .031$ ). This suggests that sounds that were common to both Spanish and English tended to be mastered earlier than English-only consonants like /ŋ, v, ʒ, z, ʃ, θ, h, dʒ, ɹ/ irrespective of how complex they were.

	Overall PCC		PCC Shared		PCC Unshared	
	M	SD	M	SD	M	SD
Bilinguals in present study	<b>72.64</b>	<b>11.18</b>	<b>78.86</b>	<b>8.56</b>	<b>61.94</b>	<b>12.31</b>
Bilinguals in FS and G’s study	72.31	12.45	77.18	11.43	62.04	18.99
Monolinguals in FS and G’s study	84.10	8.20	87.75	6.60	75.39	16.67
Shared sounds: /p/, /b/, /t/, /d/, /k/, /g/, /m/, /n/, /f/, /s/, /ð/ or [ð], /tʃ/, /l/, /w/, /j/ Unshared sounds: /ŋ/, /v/, /ʒ/, /z/, /ʃ/, /θ/, /h/, /dʒ/, /ɹ/ (English); /ɲ/, [β], /x/, [ç], /r/, /ç/ (Spanish).						

**Table 4.** Means and standard deviations for overall percentage of consonants correct (PCC) and PCC for shared and unshared sounds between Spanish and English for the bilingual subjects and the bilingual and monolingual children in Fabiano-Smith and Goldstein (2010b).

## 4. DISCUSSION, IMPLICATIONS, AND LIMITATIONS

### 4.1. Discussion

The purpose of this study was to examine the English phonological skills of ten Spanish L1/English L2 bilingual preschoolers from Mexican-origin families residing in Los Angeles, California. We assessed (1) the complexity of their English phonetic inventories, (2) the overall accuracy of their consonant production, (3) the accuracy of production of early-, middle-, and late- developing consonant sounds, and (4) the instances of interaction between phonological systems, as evidenced by transfer from Spanish and accuracy rates with shared and unshared sounds. We compared our results to those obtained in previous studies of Spanish-English bilingual preschoolers raised in Pennsylvania and Texas to assess whether bilinguals from California display similar English phonological skills.

The first analysis indicated that most participants displayed English inventories with the highest complexity levels for feature use (Dinnsen et al., 1990). These inventories, based on the sounds the children produced *accurately* in a minimum of two words, were, in complexity, commensurate to the inventories of same-age monolinguals and of bilingual preschoolers from other communities (Fabiano-Smith and Barlow, 2010; Gildersleeve-Neumann et al., 2008). The similarities in segmental inventories between our participants and the monolingual peers described in other studies suggest that these DLLs had age-appropriate speech production skills as well as phonological knowledge for English consonants. As found in previous research, our results underscore the relatively fast development of the English inventory by children learning English as an L2 during the preschool years and the similarity of this development among children from different bilingual communities.

When accuracy levels were examined, we found that the participants displayed a gradual decrease in accuracy from early-, to middle-, to late-developing English sounds, similar to the monolinguals and bilinguals in other studies (Fabiano-Smith and Goldstein, 2010a). These results mirror the findings on phonetic inventories, showing that DLLs acquire L2 sounds that are easier to produce (i.e., stops, nasals, and glides) before L2 sounds that are more complex (e.g., liquids, fricatives, and affricates). These results also provide support for Fabiano-Smith and Goldstein's (2010a) proposal to assess bilingual children's phonological skills using the accuracy of production

of EML sounds. Compared to the broader measure of overall consonant accuracy, the EML construct reveals that DLLs acquire L2 phonology in a simple-to-complex fashion as in L1.

Interestingly, our participants, who were from Mexican-origin Latino families living in Los Angeles, showed consonant accuracy rates that were very similar to those found among Latino Spanish-English DLLs speaking Puerto Rican and Dominican Spanish and living in Pennsylvania and Texas. Similar to those children, our subjects' accuracy rates tended to be lower than those of same-age monolinguals, providing additional evidence for deceleration in bilingual phonological development during the preschool years. Taken together, these findings suggest that Spanish-English bilingual preschoolers from *different communities* across the U.S. and speaking *different Spanish varieties* develop some English phonological skills (i.e. accuracy of consonant production) at a slower rate than monolinguals, while developing others (i.e. phonetic inventories) at a comparable rate. The similarity between our findings and those of previous studies underscores that this heterogeneous development may be typical in Spanish-English bilingual phonological development.

How DLLs show both a comparable and a slower rate of development of phonological skills compared to monolinguals is still a matter of debate. It has been speculated that interaction between the two phonological systems might aid in the acquisition of some phonological skills while at the same time delaying the development of other phonological skills (Anderson, 2004; Fabiano-Smith and Barlow, 2010; Fabiano-Smith and Goldstein, 2010b). In the case of phonetic inventories, DLLs may be using L1 sounds and patterns to build the L2 phonological system, in a process that has been called *phonological bootstrapping* (Fabiano-Smith and Barlow, 2010). This process would be especially felicitous in the case of phonetically similar sounds (for example, stops) which could be articulated in a similar manner in both languages without sacrificing meaning (because the essential features of phonemes – like voicing and place and manner of articulation – would be preserved). Fabiano-Smith and Barlow (2010) and Fabiano-Smith and Goldstein (2010b) define this type of interaction as a variant of acceleration. The authors argue that, thanks to phonological bootstrapping, DLLs can build inventories as complex as those of monolinguals and can maintain similar rates of acquisition as their monolingual counterparts in both languages, thus demonstrating phonological skills that are within the typical range for their chronological age. Since this study only examined the

children's phonological skills in one language, its findings cannot provide evidence for "a similar level of phonological skill (i.e., a type of acceleration) in both languages in comparison to monolinguals" (Fabiano-Smith and Goldstein, 2010b: 27). However, our participants did transfer sounds from L1 to L2 and they did acquire complex English inventories in a relatively short period of time, suggesting that phonological bootstrapping – and some form of acceleration – might have been at play.

It is possible that this type of acceleration exists only when learning languages with a high degree of phonetic similarity or overlapping phonological properties. As shown in this study, the English consonants that were present in most inventories and that were reproduced with the highest accuracy were sounds common to both Spanish and English, that is, sounds that were phonetically similar between two languages. Unshared sounds – or sounds that existed only in English – were not only less present in the children's inventories but they were also often substituted with other phones, resulting in production errors. Thus, while interaction between phonological systems bootstrapped the development of L2 inventories (positive transfer), it also resulted in negative transfer in the case of English-only sounds, lowering overall accuracy rates. These results point to the asynchronous development of different L2 phonological skills in DLLs, suggesting that inventories are mastered before accuracy of production. At the same time, our findings disconfirm Flege's proposal that L2 sounds that are dissimilar from L1 are easier to acquire than similar sounds because they are perceived differently by the learner. It is possible that Flege's proposal accounts for speech sound acquisition among adult L2 learners, for whom L1 categories are more robust and firmly established than in young learners. In the case of early bilinguals, however, the results of this study suggest that sound dissimilarity brings about negative transfer, delaying rather than helping L2 sound acquisition, as proposed by Fabiano-Smith and colleagues and Goldstein and Bunta (2012).

Whether sounds must be phonemic in both languages to aid L2 phonological acquisition remains a matter of debate. In this study, the sounds that were coded as 'shared' (or phonetically similar) and were bootstrapped from L1 to L2 were primarily phonemic in both Spanish and English. Only one 'similar' consonant – the voiced interdental fricative – was phonemic in English but allophonic in Spanish. Results for this sound were contradictory. Most English inventories displayed /ð/, suggesting that sound frequency in L1 was sufficient for phonological bootstrapping to occur during the build-

up of the L2 phonetic inventory, irrespective of whether the sound was a phoneme or an allophone (Fabiano-Smith and Goldstein, 2010a). However, although the children successfully transferred the voiced interdental fricative from Spanish to English to build the English phonetic inventory, their overall accuracy of production with this sound was low compared to monolinguals, suggesting that phonetic similarity may not be enough to bootstrap accuracy rates. When it came to PCC, the voiced interdental fricative behaved indeed like sounds coded as ‘unshared’ between Spanish and English, which were consistently reproduced with lower accuracy than in the case of monolinguals. Whether this is due to the phonological dissimilarity between English phonemic /ð/ and Spanish allophonic [ð] is arguable. Clearly, more studies are needed to examine whether positive transfer, negative transfer or both might occur while acquiring L2 sounds that are phonemic in one language but allophonic in the other.

It has been speculated that negative transfer may be most noted at the initial stages of L2 learning, when children already possess articulatory skills in one language that can be mapped to the L2 (Anderson, 2004). As the learner is more exposed to the L2, however, transfer from L1 should decrease and “L2-specific principles [should] exert a much larger influence on the development of the L2 phonology” (Anderson, 2004: 200). Lower accuracy rates in DLLs do appear indeed to be a temporary phenomenon characteristic of the early stages of English L2 development. Goldstein and Washington (2001) found that four-and-a-half-year-old Spanish/English Latino bilinguals did not differ significantly from monolingual peers in their accuracy rates; similar results were also found by Goldstein, Fabiano and Washington (2005) for five-year-old Spanish/English bilinguals with varying degrees of proficiency in each language. In addition, Goldstein and Bunta (2012), who compared the phonological skills of Spanish-English bilingual 6-year-olds to those of monolinguals with the same levels of language use and proficiency, found that at such high levels of bilingual proficiency, bilinguals outperformed their monolingual peers on many phonological measures. It appears then that, during the preschool years, Spanish L1/English L2 DLLs from different bilingual communities might be less accurate, yet still within the typical range, compared to monolingual peers in the production of English sounds. This difference, however, should cease to exist within a year or two and school-aged bilingual children who are as fluent in each language as their monolingual peers might even show an advantage.



## 4.2. Implications

The results of our study together with those of similar studies have important implications for educators and speech-language pathologists. First, in order to avoid under- and over-diagnosis of speech sound disorders among DLLs, a variety of measures should be used during phonological evaluation, especially during the preschool years, as some phonological skills might appear typical whereas others may be somewhat delayed. With regard to phonetic inventories, Spanish L1/English L2 preschoolers should not be expected to display sparser English inventories due to bilingual learning. Rather, these children should develop English inventories within a relatively short period of time and with the same complex features as those of monolingual peers. At the same time, however, instances of transfer from Spanish to English should be expected, in the form of specific error patterns that derive from L1 (for example, the use of [β] in place of /v/, of /r/, /r/, or its variants instead of /ɹ/, and the use of /x/ in place of /h/) and of more difficulty with English-only sounds (/v, z, θ/). Educators and speech-language pathologists should be aware that phonological transfer from L1 to English may be typical in bilingual development (and not an indicator of clinical speech disorder or delay) and should create instructional and intervention activities that specifically target difficult sounds. For speech-language pathologists, this might mean having children build extensive practice with English sounds that do not exist in Spanish and working on the treatment of articulation errors that might originate from L1 as discussed above. For educators working with Spanish-English bilingual preschoolers, it might mean focusing on phonemic awareness activities that specifically foster the hearing, identification, and manipulation of English-specific sounds present in words like *vase*, *think*, *zebra* etc. and perhaps spending less time on building awareness of the more 'obvious' sounds that are common to Spanish as well.

Accuracy rates among bilinguals that are slightly below average monolingual rates should also not be seen as an indicator of clinical speech delay or disorder. As pointed out by Fabiano-Smith and Goldstein (2010b), these rates are still within the typical range for English-speaking children at this age. In our study, for instance, seven of the ten children had overall consonant accuracy greater than 70%, rates that classify them in the typical range (Shriberg and Kwiatkowski, 1982). Most importantly, as argued by Fabiano-Smith and Barlow (2010), slightly lower accuracy *together with* inventories within the norm may be typical in bilingual phonological development during the preschool years. This means that a preschooler

with significantly lower accuracy rates and a sparser English inventory than the average monolingual might need speech and language services whereas a peer with slightly lower accuracy rates but a typical inventory might not. While additional research with larger numbers of children is necessary to further confirm these suggestions, this and previous studies provide important insights for professionals working with the growing dual-language population.

### 4.3. Limitations and future directions

This study has several limitations that should be considered when conducting future studies on phonological skills in DLLs. The major limitation of the current study is the small number of participants. Clearly, we cannot determine typical English skills for Spanish L1/English L2 bilingual children on the basis of ten or even twenty subjects. Future studies should include larger, homogeneous samples of children from different Spanish-English dialectal communities (Mexican-descent children living in California, Puerto Rican-descent children living in New York, etc.) and, given the heterogeneous nature of DLLs, children with different proficiency levels in L1 and L2. Despite the small number of participants, our study validated the important, yet preliminary findings of Fabiano-Smith and colleagues' and Gildersleeve-Neumann et al.'s (2008) studies, suggesting that Spanish-English bilingual children speaking different dialects (e.g., Mexican vs. Puerto Rican Spanish) and living in diverse areas of the U.S. (e.g., Pennsylvania, Texas, and California) display very similar patterns in the development of English phonological skills. Future research must build from and extend these studies by increasing the number of subjects, thereby contributing to our current understanding of typical Spanish-English phonological development.

Another limitation is that we only looked at English phonological skills, "allow[ing] a window into only one of the two languages of the bilingual child" (Gildersleeve-Neumann et al., 2008: 323). Knowledge of how Spanish develops in these children can help explain phonological patterns found in English and can ultimately provide important insights into the mechanisms and processes behind bilingual phonological development. In addition, this study did not include a monolingual comparison group and did not assess the children's phonological skills longitudinally. Clearly, future studies of phonological development among DLLs must examine how phonological skills in DLLs compare to those of same-background monolinguals and how they change from the initial to the later stages of L1 and L2 development.



The final limitation of this study – or perhaps an alternative explanation for our findings – is that the features of the English dialect spoken by the children were reflected in the results. As shown above, for instance, the children had particularly low accuracy rates with specific sounds, in particular /v, ð, z, θ/. Errors included substituting /v/ with the bilabial Spanish fricative [β], devoicing /z/, and stopping /ð/ and /θ/. It was argued that these sounds were difficult for the children because, with the exception of the voiced inter-dental fricative, which exists at the allophonic level in Spanish, these sounds were unique to English. Interestingly, however, these same processes are typical features of Chicano English, a non-standard English variety influenced by Spanish spoken as a native dialect both by bilingual and monolingual speakers of Mexican ethnic origin in California and the Southwest (Fought, 2003). Thus the possibility exists that what we coded as “errors” by our participants (and hence the lower accuracy rates) were not the result of phonological development but rather were phonological processes produced even by *monolingual* English-speaking children from the same Mexican-American community. Fought (2003) and Santa Ana and Bayley (2004) clearly describe the substitution of English inter-dental fricatives by apico-dental stops as a typical phonological feature of Chicano English, even in the case of speakers who do not speak Spanish. Penfield (1984) also provides a detailed account of /v/ stopping and /z/ devoicing as characteristics of Chicano English. Interestingly, even Gildersleeve-Neumann et al. (2008), who examined the phonological skills of Spanish-English DLLs in Texas, found that this sort of devoicing did not decrease over time but continued in frequency over a one-year period for the Spanish-English bilingual preschoolers in their study. It is thus possible that the differences that we observed between our participants and English-speaking monolinguals were due to dialectal differences.

Informal observations of English monolingual children from the same community as our subjects provide some support for this hypothesis. The Mexican-American community in Los Angeles is indeed very strong and in certain areas – such as those where the children lived – the majority of the population is of Mexican-origin. The high level of immigration from Mexico during the last century and the relative insularity of this ethnic group with respect to the Anglo community have allowed the development of a linguistically separated community whose members are not always bilingual but at times monolingual in English. Preschool teachers in these areas tend to be members of the same community (Whitebook, Kipnis, and Bellm, 2008),

and the majority of these teachers also speak Chicano English. In sum, we cannot rule out the possibility that our results were in part due to the dialect spoken by the children and the adults around them.

Other studies have documented the influence of the individual's speech community on phonological production in English for Spanish L1 speakers. Wolfram (1971), for instance, found that native Spanish speakers' English pronunciation was influenced by the English variety spoken by the members of the same speech community, in this case, African American English. Other researchers further discuss the difficulties of studying language development in populations that do not speak the standard variety. Stockman (2010), for instance, argues that identifying speakers of African American English with speech-language delays can be difficult because some of their typical dialect patterns (e.g., [f]-for-/θ/ substitutions and copular verb absence) can resemble those used by typical speakers of Standard American English with speech-language delay. Goldstein and Gildersleeve-Neumann (2012) argue that similar difficulties apply when studying speakers of non-standard Spanish and English varieties because dialect differences might easily be interpreted as errors. In a study of Spanish phonological development with typically-developing children and children with speech sound disorders, Goldstein and Iglesias (2001) found indeed that the number of consonant errors and phonological processes decreased when features of Puerto Rican Spanish were taken into account. Similarly, Shriberg and Kwiatkowski (1982) found that overall consonant accuracy increased after dialectal features were accounted for. While taking into account dialectal features did not change the classification of any child from being phonologically disordered to typically developing, this research points to the importance of recognizing the influence of dialect on L1 and L2 production.

Interestingly, in their recent studies, Fabiano-Smith and colleagues argued that their participants spoke the English dialect common to the northeastern United States (a relatively standard variety) and only [n]-for-/ŋ/ substitutions were not coded as errors. Since the goal of the present investigation was to compare the phonological skills of those children to those of Spanish-English bilingual children from Los Angeles, we followed the same methodology. However, while it is possible that the participants of Fabiano-Smith and colleagues' studies were learning an English variety that was not influenced by Spanish (after all, they lived in Pennsylvania, a region geographically distant from Spanish-speaking Mexico and Central

America), we raise the possibility that the situation was different for the children in our study. Southern California Mexican-Americans form a strong ethno-linguistic group, with its own dialect. It is possible that these speakers might learn Spanish-influenced Chicano English as their native language, irrespective of whether they are bilingual or monolingual English speakers. Clearly, any future study of this population should take into account the dialectal features of both Mexican Spanish and Chicano English and include monolingual children from the same speech community to provide the most accurate picture of the developing phonological skills of Spanish-English bilingual children.

## REFERENCES

- Amastae, Jon (1995): "Variable spirantization: Constraint weighting in three dialects", *Hispanic Linguistics* 6/7, pp. 267–285.
- Anderson, Raquel (2004): "Phonological acquisition in preschoolers learning a second language via immersion: A longitudinal study", *Clinical Linguistics and Phonetics* 18(3), pp. 183–210.  
<http://dx.doi.org/10.1080/0269920042000193571>
- Barlow, Jessica (2003a): "The stop-spirant alternation in Spanish: Converging evidence for a fortition account", *Southwest Journal of Linguistics* 22(1), pp. 51–86.
- (2003b): *Assessment of English Phonology*. San Diego, CA: Phonological Typologies Laboratory, School of Speech, Language, and Hearing Sciences, San Diego State University.
- Children Now (2014): 2014 California Children's Report Card. Oakland, CA: Author.  
[http://www.childrennow.org/uploads/documents/2014\\_CA\\_Childrens\\_Report\\_Card.pdf](http://www.childrennow.org/uploads/documents/2014_CA_Childrens_Report_Card.pdf).
- Dinnsen, Daniel A., Steven B. Chin, Mary Elbertand Thomas Powell (1990): "Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics", *Journal of Speech and Hearing Research* 33, pp. 28–37.  
<http://dx.doi.org/10.1044/jshr.3301.28>
- Eilers, Rebecca E., D. Kimbrough Oller and Carmen R. Benito-García (1984): "The acquisition of voicing contrasts in Spanish and English learning infants and children: A longitudinal study", *Journal of Child Language* 2, pp. 313–336.

- Fabiano-Smith, Leah and Jessica Barlow (2010): "Interaction in bilingual phonological acquisition: Evidence from phonetic inventories", *The International Journal of Bilingual Education and Bilingualism* 31(1), pp. 1-17.
- and Brian A. Goldstein (2010a): "Early-, middle-, and late-developing sounds in monolingual and bilingual children: An exploratory investigation", *American Journal of Speech-Language Pathology* 19(1), pp. 66-77.  
[http://dx.doi.org/10.1044/1058-0360\(2009/08-0036\)](http://dx.doi.org/10.1044/1058-0360(2009/08-0036))
- and Brian A. Goldstein (2010b): "Phonological acquisition in bilingual Spanish-English speaking children", *Journal of Speech, Language, and Hearing Research* 53, pp. 160-178.  
[http://dx.doi.org/10.1044/1092-4388\(2009/07-0064\)](http://dx.doi.org/10.1044/1092-4388(2009/07-0064))
- Flege, James E. (1981): "The phonological basis of foreign accent: A hypothesis", *TESOL Quarterly* 15(4), pp. 443-455.  
<http://dx.doi.org/10.2307/3586485>
- (1995): "Second language speech learning: Theory, findings and problems", in Winifred Strange (ed.): *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York Press, pp. 233-277.
- Fought, Carmen (2003): *Chicano English in Context*. New York: Palgrave/MacMillan Publishers.
- Genesee, Fred and Elena Nicoladis (2007): "Bilingual first language acquisition", in Erika Hoff and Marilyn Shatz (eds.): *Blackwell Handbook of Language Development*. Malden, MA: Wiley-Blackwell, pp. 324-342.  
<http://dx.doi.org/10.1002/9780470757833.ch16>
- Gildersleeve-Neumann, Christina E.; Ellen Kester, Barbara L. Davis and Elizabeth D. Peña (2008): "English speech sound development in preschool-aged children from bilingual Spanish-English environments", *Language, Speech, and Hearing Services in Schools* 39, pp. 314-328.  
[http://dx.doi.org/10.1044/0161-1461\(2008/030\)](http://dx.doi.org/10.1044/0161-1461(2008/030))
- Goldstein, Brian A. and Ferenc Bunta (2012): "Positive and negative transfer in the phonological systems of bilingual speakers", *International Journal of Bilingualism* 16(4), pp. 388-401.  
<http://dx.doi.org/10.1177/1367006911425817>

- , Leah Fabiano and Patricia S. Washington (2005): “Phonological skills in predominantly English, predominantly Spanish, and Spanish-English bilingual children”, *Language, Speech, and Hearing Services in Schools* 36, pp. 201–218.  
[http://dx.doi.org/10.1044/0161-1461\(2005/021\)](http://dx.doi.org/10.1044/0161-1461(2005/021))
  - and Christina E. Gildersleeve-Neumann (2012): “Phonological development and disorders”, in Brian A. Goldstein (ed.): *Bilingual Language Development and Disorders in Spanish-English Speakers* (2nd ed.). Baltimore, MD: Brookes, pp. 285-309.
  - and Aquiles Iglesias (2001): “The effect of dialect on phonological analysis: Evidence from Spanish-speaking children”, *American Journal of Speech-Language Pathology* 10, pp. 394-406.  
[http://dx.doi.org/10.1044/1058-0360\(2001/034\)](http://dx.doi.org/10.1044/1058-0360(2001/034))
  - and Patricia S. Washington (2001): “An initial investigation of phonological patterns in 4-year-old typically developing Spanish-English bilingual children”, *Language, Speech, and Hearing Services in Schools* 32, pp. 153–164.  
[http://dx.doi.org/10.1044/0161-1461\(2001/014\)](http://dx.doi.org/10.1044/0161-1461(2001/014))
- Guitart, Jorge (1996): “Spanish in contact with itself and the phonological characterization of conservative and radical styles”, in Ana Roca and John B. Jensen (eds.): *Spanish in Contact: Issues in Bilingualism*. Somerville, MA: Cascadilla Press, pp. 151-157.
- Gutiérrez-Clellen, Vera (1996): “Language diversity: Implications for assessment”, in Kevin N. Cole, Philip S. Dale, Donna J. Thal (eds.): *Assessment of Communication and Language*. Baltimore, MD: Brookes, vol. 6, pp. 29-56.
- Lidz, Carol S. and Elizabeth D. Peña (1996). “Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children”, *Language, Speech, and Hearing Services in Schools* 27, pp. 367–372.  
<http://dx.doi.org/10.1044/0161-1461.2704.367>
- Lozano, María del Carmen (1979): *Stop and Spirant Alternation: Fortition and Spirantization Processes in Spanish Phonology*. Bloomington: Indiana University Linguistics Club.

- Macken, Marlys A. and D. Barton (1980): "The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants", *Journal of Child Language* 7, pp. 433-458.  
<http://dx.doi.org/10.1017/S0305000900007029>  
<http://dx.doi.org/10.1017/S0305000900002774>
- Mayr, Robert and Simona Montanari (in press): "Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast", *Journal of Child Language*.
- Miller, Jon F. and Aquiles Iglesias (2006): *Systematic Analysis of Language Transcripts (SALT), Version 9*. University of Wisconsin, Madison.
- National Assessment of Educational Progress (NAEP) (2011): *Fast Facts*. Washington, DC: Author.  
<http://nces.ed.gov/fastfacts/display.asp?id=147>.
- National Center for Family Literacy (2008): *Developing Early Literacy: Report of the National Early Literacy Panel*. Washington, DC: Author.  
<http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>.
- National Clearinghouse for English Language Acquisition (NCELA) (2011): *Key Demographics and Practice Recommendations for Young English Learners*. Silver Spring, MD: Author.  
<http://www.ncela.us/files/uploads/9/EarlyChildhoodShortReport.pdf>.
- National Task Force on Early Childhood Education for Hispanics (2007): *Para Nuestros Niños: Expanding and Improving Early Education for Hispanics. Executive report*. Tempe, AZ: Author.
- Paradis, Johanne and Fred Genesee (1996): "Syntactic acquisition in bilingual children: Autonomous or interdependent?", *Studies in Second Language Acquisition* 18, pp. 1-25.  
<http://dx.doi.org/10.1017/S0272263100014662>
- , Fred Genesee and Martha Crago (2011): *Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning* (2nd ed.). Baltimore, MD: Brookes.
- Penfield, Joyce (1984): "The vernacular base of literacy in Chicano English", in Jacob Ornstein-Galicia (ed.): *Form and Function in Chicano English*. Rowley, MA: Newbury House, pp. 71-84.



- Perissinotto, Giorgio (1975): *Fonología del Español Hablado en la Ciudad de México: Ensayo de un Método Sociolingüístico*. Translated by Raúl Ávila. México: Colegio de México.
- Santa Ana, Otto and Robert Bayley (2004): “Chicano English phonology”, in Edgar Werner Schneider, Bernd Kortmann, Kate Burridge, Rajend Mesthrie and Clive Upton (eds.): *A Handbook of Varieties of English: Phonology*. Berlin: Mouton de Gruyter, pp. 407–424.
- Shriberg, Lawrence D. (1993): “Four new speech and voice-prosody measures for genetics research and other studies in developmental phonological disorders”, *Journal of Speech, Language, and Hearing Research* 36, pp. 105–140.  
<http://dx.doi.org/10.1044/jshr.3601.105>
- , Diane Austin, Barbara A. Lewis, Jane L. McSweeney and David L. Wilson (1997): “The percentage of consonants correct (PCC) metric: Extensions and reliability data”, *Journal of Speech, Language, and Hearing Research* 40, pp. 708–722.  
<http://dx.doi.org/10.1044/jslhr.4004.708>
- and Joan Kwiatkowski (1982): “Phonological disorders III: A procedure for assessing severity of involvement”, *Journal of Speech and Hearing Disorders* 47, pp. 256–270.  
<http://dx.doi.org/10.1044/jshd.4703.256>
- Stockman, Ida J. (2010): “A review of developmental and applied language research on African American children: From a deficit to a different perspective in dialect differences”, *Language, Speech, and Hearing Services in School* 41, pp. 23–38.  
[http://dx.doi.org/10.1044/0161-1461\(2009/08-0086\)](http://dx.doi.org/10.1044/0161-1461(2009/08-0086))
- Stoel-Gammon, Carol (1987): “Phonological skills of 2-year-olds”, *Language, Speech, and Hearing Services in Schools* 18, pp. 323–329.  
<http://dx.doi.org/10.1044/0161-1461.1804.323>
- (2011): “Relationships between lexical and phonological development in young children” *Journal of Child Language* 38(1), pp. 1–34.  
<http://dx.doi.org/10.1017/S0305000910000425>
- Whitebook, Marcy, Fran Kipnis and Dan Bellm (2008): *Diversity and Stratification in California’s Early Care and Education Workforce*. Berkeley, CA: Center for Child Care Employment, UC Berkeley.



- Whitley, Melvin S. (2002): *Spanish/English Contrasts: A Course in Spanish Linguistics* (2nd ed.). Washington, DC: Georgetown University Press.
- Wolfram, Walt. (1971). *Overlapping Influence and Linguistic Assimilation in Second Generation Puerto Rican English*. Paper presented at the meeting of the American Anthropological Association, New York, NY.



## **Aspectual Distinctions in Spanish: A Case Study of Transfer in Second-Language Learners**

| Ana Sánchez-Muñoz  
| *California State University, Northridge*

**Abstract:** Languages that are typologically similar may show positive transfer effects but, conversely, negative transfer effects can be expected in areas in which the two languages differ (Odlin 1989: 26). However, the situation may not be so clear in areas that are only partially similar in the native and target language. This study looks into the acquisition of one of these types of constructions in Spanish and English compared to structures in which the two languages differ.

The particular focus is on the acquisition of two Spanish grammatical structures by intermediate and advanced English-speaking learners: the progressive constructions and preterit-imperfect distinction. The first of these structures is similar in form in both languages and partially similar, but not identical, in conditions of use. On the other hand, the preterit-perfective distinction constitutes an area of difference between Spanish and English since in Spanish the [ $\pm$ perfective] distinction is manifested with overt tense morphology.

**Keywords:** Spanish, Second-Language, Transfer, Acquisition, Aspect.



## 1. INTRODUCTION

As early as the 1940s, the study on transfer started capturing the interest of linguists investigating language classification and change. It was not until the 1970s and 1980s that the phenomenon of transfer began to be seriously considered in the research of child language acquisition and second language acquisition (Ringbom and Palmberg 1976; Schachter and Rutherford 1979; Apple 1984). From then on, the issue of language transfer, that is, the influence that a language has on another, has become a major question addressed in both fields: bilingual/multilingual language acquisition and second language acquisition.

The present paper explores the acquisition of structures that constitute areas of difference between L1 and L2 compared to the acquisition of similar structures in both languages. This study concentrates on the acquisition of two Spanish grammatical structures by intermediate and advanced English-speaking learners: 1) the progressive constructions and 2) the acquisition of the aspectual distinction preterit-imperfect. The first of these structures has similar form in both languages and partially similar, but not identical, conditions of use. The second structure is different in Spanish and in English.

In second language acquisition, languages that are typologically similar, such as Spanish and English there may be (and there normally is) evidence of positive transfer; that is, cases in which the acquisition of the second language is facilitated by the similarities between the native and the target language (Odlin 1989: 36). However, what happens in structures that are only partially (not completely) similar? Can we expect positive or negative transfer in these cases? If there is transfer in these two types of structures (i.e. similar and different structures in L1 and L2), it is important to examine whether the effects of transfer are the same for the structures examined or if, on the other hand, there are noticeable differences, either quantitative or qualitative, between the results obtained for similar vs. different constructions.

With these considerations in mind, this paper will try to address questions such as whether there is any transfer effect in the two types of structures mentioned above, namely structures that are completely different and structures that share some similarities in the native and the target language. Also, if there is transfer, how different (if any) the effects of transfer are in those two types of structures. And, finally, we will try to determine

whether the evidence obtained from the data is enough to conclude that there are different effects of transfer (qualitative vs. quantitative) in Spanish second language acquisition depending on the degree of similarity or difference of the structures.

## 2. LANGUAGE TRANSFER: DEFINITIONS AND PROBLEMS OF STUDY

There are many theoretical and practical problems that attend the study of transfer. It is beyond the scope of this study to examine them all, but the most basic problem of terminology and definition will be considered.

The first difficulty is how to define the term transfer. Different scholars refer to this linguistic phenomenon in somewhat different ways as the long-standing use of transfer has led to differences of opinion. In this section, some definitions of the term transfer will be presented along with critiques for those definitions.

First, some scholars have seen transfer as simply a consequence of habit formation in the behaviorist sense. However, this is not the case since the behaviorist notion of transfer is quite different from the notion of native language influence (Carroll 1968): The latter often implies the extinction of earlier habits, whereas the former need not lead to any replacement of the learner's primary language.

Second, transfer has been defined as interference from a language into another language. Weinreich (1953/1968) used the term interference to cover any case of transfer. However, more recent research differentiates interference or negative transfer from positive transfer, which refers to the facilitation influence of the similarities between two languages.

Third, transfer is not simply falling back on the native language. This is specifically relevant for second language acquisition. However, a similar claim can be made of bilingual child acquisition on the relationship between the child's dominant and weak languages; in this case, transfer would not simply be a "falling back" to the dominant language as there are other factors such as the type of input.

With these considerations in mind, we could take a broad definition of transfer as the influence that a language has on another language. In this paper, we are specifically concerned with the influence of the learner's L1 into the system of the L2.

### 3. THE PROGRESSIVE CONSTRUCTION AND THE PRETERIT-IMPERFECT DISTINCTION IN SPANISH

This study is concerned with the acquisition of Spanish as a second language and the possibility of transfer in structures with different levels of similarity between the native (English) and the target language (Spanish). Two structures are examined in this paper: 1) the progressive constructions, which have similar –but not identical– conditions of use in English and Spanish. 2) The acquisition of the aspectual distinction preterit-imperfect. Let us next briefly describe the form and uses of the progressive constructions and the preterit-imperfect distinction in Spanish in order to identify the differences in form and/or function between Spanish and English.

#### 3.1. The Spanish progressive construction

The progressive construction has similar form in both languages and partially similar, but not identical, conditions of use. Spanish and English have progressive constructions that are parallel in form and partially similar in function. Both languages coincide in the use of the progressive form to refer to action in progress at the time of speaking (example 1), while the simple present form indicates more general situations (example 2) (Comrie 1976: 33; Solé and Solé 1977: 42):

- (1) Mira, está lloviendo  
look is-it raining  
'Look, it is raining'
- (2) Llueve más en invierno que en verano  
rains-it more in winter than in summer  
'It rains more in winter than in summer'

In these cases, we may expect that the similarities between the languages facilitate the acquisition of this structure in Spanish.

However, we may also expect some negative transfer effects given that the Spanish system of progressive constructions does not exactly correspond to the English one. According to some grammarians the differences are the following (e.g. Solé and Solé 1977; Butt and Benjamin 1994:230-232):

First, Spanish can choose between the simple (as in 3a) and the progressive form (as in 3b) to refer to present ongoing activity whereas English can only use the progressive form in this context (4b). Note that in English the simple present is ungrammatical (4a).

- (3) a. Ana lee  
 Ana read-she  
 'Ana reads'  
 b. Ana está leyendo  
 Ana is-she reading  
 'Ana is reading'
- (4) a. ¿Qué hace Juan ahí afuera?  
 what does-he Juan there outside  
 '\*What does John do outside?'  
 b. ¿Qué está haciendo Juan ahí afuera?  
 what is-he doing Juan there outside  
 'What is Juan doing outside?'

The effect of transfer in this area of syntax would lead to the increase of the frequency of use of a form, the progressive present, in detriment of a variant form, the simple present.

The second difference is that in English the progressive doubles as a future tense whereas Spanish uses the simple present or the future (either inflected or periphrastic):

- (5) Llegamos (\*estamos llegando) mañana  
 arrive-we are-we arriving tomorrow  
 'We're arriving tomorrow'

Third, the continuous form in Spanish is not used to describe states rather than actions whereas English often allows the progressive form for states as illustrated in 6 and 7:

- (6) Pareces (\*estás pareciendo) cansada hoy  
 look-you are-you looking tired today  
 'You're looking tired today'
- (7) ¡Qué bien huele (\*está oliendo) la madreselva hoy!  
 how well smell-it is-it smelling the honeysuckle today  
 'Isn't the honeysuckle smelling good today!'

Cases of transfer in utterances such as 6 and 7 would indicate the extension of the native language (i.e. English) meaning and function. That is to say, the incorporation of the meaning of a form (the progressive form) from the L1 English into the meaning of the same form in the L2 Spanish (Weinreich 1974; Anderson 1984).

In cases such as 5-7 above, it is possible to think that the effects of transfer will be qualitative, that is, language learners will be likely to produce the progressive constructions in cases in which Spanish does not allow the progressive. However, based on the findings in first and second language acquisition, we expect that the progressive marking is not incorrectly over-generalized to statives (Li and Shirai 2000). Accordingly, it is hypothesized that any effect of transfer between L1 English and L2 Spanish will be manifested only in quantitative terms.

### 3.2. The Spanish preterit and imperfect distinction

The second type of Spanish structures considered in this paper is one that differs in form and/or function from English such as the distinction between preterit/imperfect. Spanish distinguishes two types of aspectual past: imperfect past, and perfect (preterit). In the former, the action expressed by the verb is viewed as unfinished (as in example 8, i.e. we were in the process of getting to the house), conversely, the latter expresses an action viewed as completed (as in example 9, i.e. we got to the house):

- (8) Caminábamos a la casa  
walk-IMP-WE to the house  
'We walked to the house'
- (9) Caminamos a la casa  
walk-PRET-WE to the house  
'We walked to the house'

Unlike English, in Spanish the [ $\pm$ perfective] distinction is manifested with overt tense morphology. In example 8, the imperfect exemplifies the imperfective value while in example 9, the preterit tense exemplifies the perfective value of the viewpoint distinction.

According to Comrie (1976), the imperfective value of the [ $\pm$ perfective] subsumes at least one more opposition: that between progressive and habitual. Both view the situation from within. Furthermore, a progressive event is by definition a dynamic event, unfolding over a period including the reference point. The habitual, on the other hand, is usually a series of finished events.

Several syntactic and semantic properties have been identified for the preterit/imperfect distinction in Spanish (Montrul and Slabakova 2000). Due to the complexity of the aspectual distinction, here we will concentrate just on one condition: the effect that the choice of the preterit or imperfect



has on the implications for the nature of the event (completed or habitual respectively). On the other hand, the English past tense interpretation depends on whether the verb phrase is an activity or an accomplishment.

- (10) *María comía en ese restaurante*  
 María eat-IMP-she in that restaurant  
 ‘María (habitually) ate in that restaurant’
- (11) *María comió en ese restaurante*  
 María eat-PRET-she in that restaurant  
 ‘María ate (once) in that restaurant’

Since the English equivalent “María ate in that restaurant” combines an activity with the past tense, the interpretation is habitual (since there are no adverbials indicating the contrary).

As a result of this contrast between Spanish and English, we may expect students to experience difficulty in the acquisition of the preterit/imperfect distinction present overtly in Spanish but not in English (Montrul and Slabakova 1999; 2000). We hypothesize that, as opposed to the typologically similar structures, transfer from the L1 will be manifested in the L2 tense/aspect domain in qualitative terms, that is, in the production of structures not present in the target language.

#### 4. HYPOTHESES

In brief, the hypothesis under investigation is that transfer will occur in two types of structures: 1) structures which are different in the native and target language and 2) structures which are similar in both languages but only partially so. Transfer may take the form of qualitative and/or quantitative influence: qualitatively, one must identify structures occurring in second language learners which are not found in the natives’ language; quantitatively, it is necessary to determine any influence in terms of SLA learners’ frequency or productivity of structures instantiated in the target language.

Two sub-hypotheses follow: As far as the progressive construction is concerned, it is hypothesized that any effect of transfer between L1 English and L2 Spanish will be manifested in quantitative terms. On the other hand, the preterit-imperfect distinction is expected to yield qualitative effects by producing ungrammaticality in the target language.

## **5. EXPERIMENT DESIGN AND METHODOLOGY**

### **5.1. Participants**

Fifty English-speaking learners of Spanish took part in the experiment, divided into two proficiency groups: 20 intermediate and 30 advanced. All the participants for this study were undergraduate students enrolled in the Department of Spanish of a large research university in Southern California. The 20 intermediate-level students were enrolled in the third semester of the Department's Spanish instruction under the Basic Language Program. They had all completed two semesters of Spanish in this university or its equivalent at a different institution. The advanced-level students were all enrolled in advanced Spanish language courses beyond the three basic courses of this university's Language Program. All the participants received 1% extra-credit in their Spanish class for their participation in this experiment. Additionally, ten Spanish native speakers served as a control group.

### **5.2. Test Instruments and procedures**

The level of proficiency of the participants was measured by a cloze test. The purpose of this test was to ensure that the level of proficiency of all the participants was the minimum necessary to complete the experimental tasks. This type of cloze test, which has been used in previous studies (Montrul and Slabakova 2000; Cabrera and Zubizarreta 2002), included three paragraphs in Spanish with a total of seventy-five blank spaces (one blank space every five words).

All students had to complete several tasks contained in a booklet. There were two types of experiment booklets according to the structure examined, progressive or preterit-imperfect. Both booklets consisted of two parts: 1) A questionnaire with biographical information as well as information about their educational and language background in order to ensure the homogeneity of the sample of speakers and data. 2) Fill-in-the-gap questionnaire designed to obtain information about the form that the students considered correct in different contexts (i.e., between the simple present and the progressive form or between the preterit and the imperfect). For those contexts where both forms could be possible in Spanish (as in examples 3 and 4), the questionnaire tested the students' preference of use, as both forms would yield grammatical answers. The fill-in-the gap questionnaire consisted of

twelve small paragraphs corresponding to different situations. There were twenty gaps total, out of which eight were gaps with different syntactic questions irrelevant for the present study to divert the participants' attention from the focus of the study.

### 5.3. Data analysis

The data obtained from the experiment were analyzed statistically using Chi-square tests. The analyses were performed with the Statistical Package for the Social Sciences (SPSS).

## 6. RESULTS

### 6.1. Proficiency Test

The results of the Cloze test for the experimental group (intermediate, I and advanced, A) are shown in Figure 1. The two groups' means are statistically significant ( $p < .001$ ).

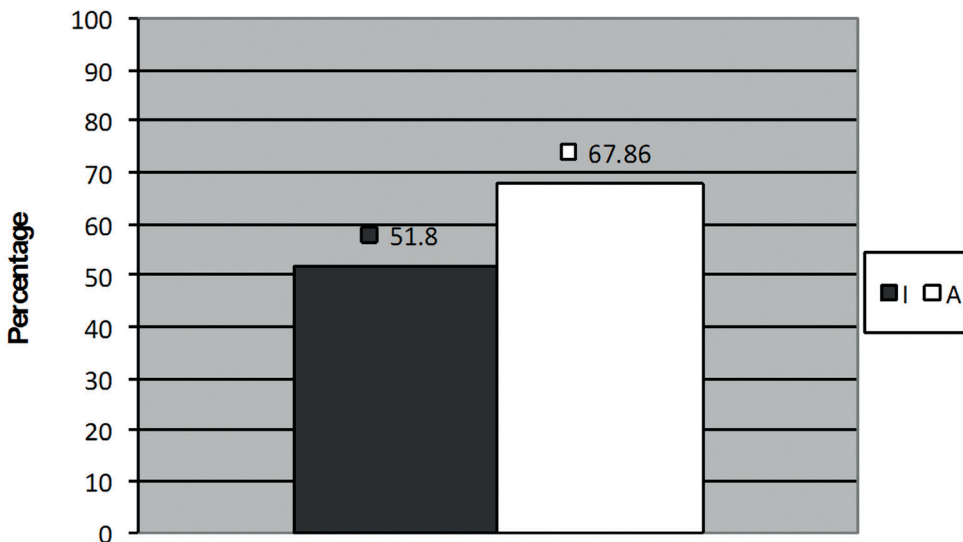


Figure 1. Mean percentage accuracy scores

## 6.2. Progressive Constructions

The experiment to test the use of the progressive constructions consisted on a fill-in-the-gap questionnaire. In the fill-in-the-gap questionnaire, the subjects had to choose which of the two supplied forms goes well in the context. The questionnaire included examples of two manifestations of transfer mentioned above: 1) The increase of use of a form (progressive) at the expense of a variant form (simple). 2) The extension of meaning of a form in a language by the incorporation of the meaning(s) of that form in another language. Examples of the tested sentences are shown in 12-14:

Progressive with present reference:

- (12) ¿Qué hace / está haciendo ahora?  
 what does-he is-he doing now  
 'What is he doing now?'

Progressive with future reference:

- (13) Vuelo(\*estoy volando) a Chicago la próxima semana.  
 fly-I am-I flying to Chicago the next week  
 'I'm flying to Chicago next week.'

Progressive with stative meaning:

- (14) Pareces (\*estás pareciendo) cansada hoy.  
 look-you are-you looking tired today  
 'You're looking tired today.'

The fill-in-the-gap questionnaire included both optional contexts for the progressive and (as in example 12) and obligatory contexts for the simple present (such as examples 13 and 14). Throughout this study, the former are called "frequency contexts" since the effects of transfer are quantitative, namely manifested in an increase of frequency of use. On the other hand, instances such as 13 and 14 are called "accuracy contexts" since the effect of transfer would be qualitative, leading to inaccurate sentences in Spanish. Figures 2 and 3 show the results of the fill-in-the-gap questionnaire for the two experimental groups (I, A) and for the control group (C) for the frequency and the accuracy condition.

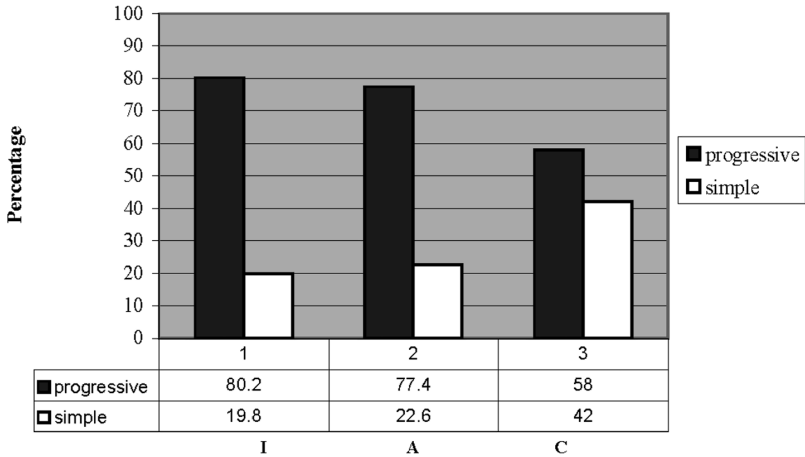


Figure 2. Progressive Constructions: Frequency Contexts

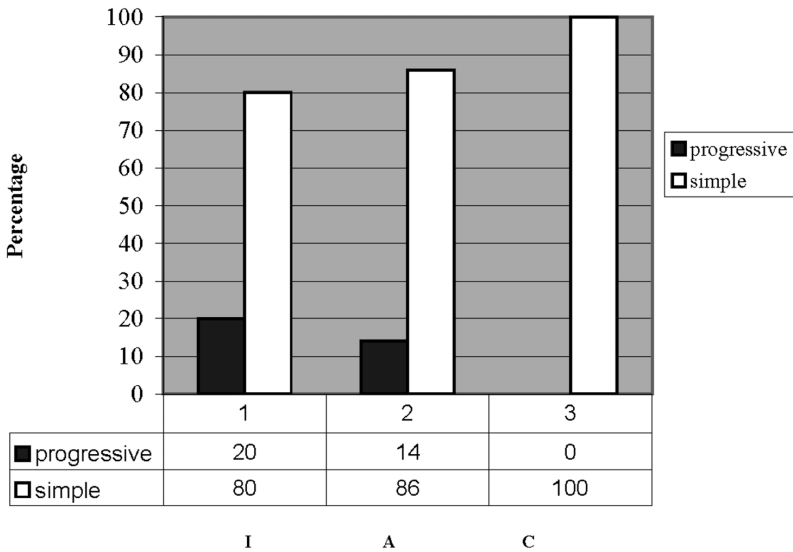


Figure 3. Progressive Constructions: Accuracy Contexts

Figure 2 shows that the two experimental groups, both I and A, chose more progressive forms (80.2% and 77.4% respectively) than the control group (58%). This difference is significant enough ( $p < .0002$ ) to argue in favor of a considerable quantitative effect due to L1 transfer. However, the

difference between the intermediate and advanced groups is very small. On the other hand, figure 3 illustrates some qualitative effects that were not expected. Unlike the control group, both the I and A group chose the progressive forms (20% and 14% respectively) for structures which did not allow the use of this form. Nevertheless, the majority of the occurrences of the progressive form were produced in contexts with future reference (such as example 11 above) and very few in stative contexts (example 12). This evidence is consistent with the claim that the progressive is not over-generalized to statives (Li and Shirai 2000; Salaberry 2000), although it is important to note that some students did so.

### 6.3. Preterit/imperfect distinction

As with the previous structure, the fill-in-the-gap questionnaire consisted of sentences in which the students had to choose between the preterit or the imperfect. The context provides the information necessary for only one form to be appropriate in each case as shown in 15 and 16:

(15) Nosotros hicimos / hacíamos un viaje a Madrid  
 el año pasado  
 we did-PRET did-IMP a trip to Madrid  
 the year past  
 'We took (PRET/IMP) a trip to Madrid last year.'

(16) Gabriel visitó / visitaba a Carmen frecuentemente.  
 Gabriel visit-PRET visit-IMP to Carmen frequently  
 'Gabriel visited (PRET/IMP) Carmen frequently.'

In example 15, the preterit is the only correct form as the habitual interpretation is excluded by the NP *el año pasado* 'last year', which specifies the one-time occurrence of the event. Conversely, the adverbial *frecuentemente* ('frequently') in example 16, provides the context for the habitual interpretation of the event thus the imperfect is the only acceptable choice. In this way, the task tests the subjects' interpretation of semantic entailments of the two tenses. The preterit denotes a one-time event in the past while the imperfect denotes a habitual event (Montrul and Slabakova 2000; 2002).

This questionnaire included 15 gaps out of which 5 were distractors. The overall accuracy on the preterit was significantly different from the imperfect for the advanced and intermediate learners. The results of this questionnaire are shown in figure 4.

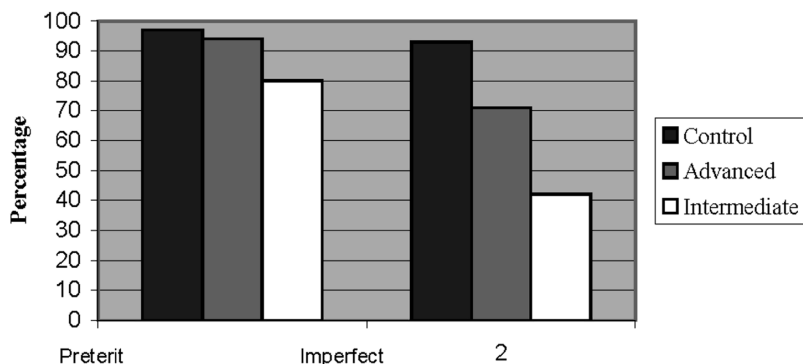


Figure 4. Mean Accuracy on Tense (Habitual vs. One-time event)

Figure 4 presents subjects' accuracy according to the verb form, either preterit or imperfect. Both the advanced and the intermediate learners were more accurate with the preterit than with the imperfect ( $p < .0002$  and  $p < .0001$ , respectively). In agreement with Montrul and Slabakova's findings (2000), these results suggest that the subjects have acquired the fact that the preterit can describe one-time event in the past, and cannot describe habitual activity. The knowledge of the imperfect semantics, however, may develop later.

## 7. DISCUSSION

The results of the two experiments reported above reveal there is a difference in the effects of transfer on second language acquisition depending on the similarity between the L1 and the L2 structures, as it was predicted.

The progressive construction, which is partially similar in Spanish and in English, is acquired faster by second language learners. Additionally, the results show that there is not a major difference between the intermediate and advanced groups. In the fill-in-the-gap questionnaire, the subjects choose more progressive forms than the control group and we can see this quantitative effect (figure 2). Furthermore, the difference is significant indicating that such effect may be due to transfer from the L1. More interestingly, the results obtained from the accuracy task in the progressive show that second language learners may be sensitive to negative transfer from native language patterns even in this type of similar structures (figure 3).

Second, the results of the tasks testing the acquisition of the Spanish past aspectual distinction show that the lack of equivalent preterit/imperfect structures in L1 may lead to the observed qualitative differences between the subjects and the control group. The intermediate and advanced groups reveal a trend in the acquisition of semantic interpretations of preterit and imperfect in the condition tested (habitual vs. one-time event). The results show that the preterit is acquired faster and subjects associate it with on-time events in the past. However, it takes more time to acquire the fact that the imperfect can describe a habitual activity as shown in the significant difference between the intermediate and advanced groups. These results are consistent with Montrul and Slabakova's (2000) conclusion that L2 learners are able to learn properties of the grammar that are more difficult due to typological differences between L1 and L2.

In conclusion, although the data analyzed may not be sufficient to arrive at any generalization, it can be said that the results obtained so far reveal a trend in the acquisition of different types of structures and on the different effects of transfer from the L1 on the L2.

Future research should also include investigation of the way in which learners interpret the L2 structures. For example, learners may produce native-like aspectual morphology but may have non-native aspectual interpretations. In other words, going beyond production data into comprehension experiments, and especially incorporating insights from theoretical research on aspectual semantics and pragmatics would be an interesting area for future research. Also, as this study has tried to show, it is very important to take into account the effect of a learner's native language on his/her acquisition of the properties of L2 structures. In the case of the acquisition of aspectual properties, the knowledge of the L1 language has been neglected. I believe it is crucial to understand what the learners know from their native English, and what exactly their learning task is in acquiring Spanish, including areas of typological contrast where they may have most difficulties such as the ones examined in the present study.





## REFERENCES

- Andersen, Roger (1983): "Transfer to Somewhere", in Susan M. Gass and Larry Selinker (eds.): *Language Transfer in Language Learning*. Rowley, MA: Newbury House, pp. 177-201.
- Apple, René (1984): *Immigrant Children Learning Dutch*. Dordrecht: Foris Publications.
- Butt, John and Carmen Benjamin (1994): *A New Reference Grammar of Modern Spanish*. Lincolnwood, IL: NTC.
- Cabrera, Mónica and María Luisa Zubizarreta (2003): "On the acquisition of Spanish causative structures by L1 speakers of English", in Juana M. Liceras, Helmut Zobl and Helen Goodluck (eds.): *Proceedings of The 2002 Generative Approaches to Second Language Acquisition (GASLA-6): L2 Links*. Somerville, Mass.: Cascadilla Press, pp. 24-33.
- Comrie, Bernard. (1976): *Aspect*. Cambridge: Cambridge University Press.
- Carroll, John B. (1968): "Contrastive Linguistics and Interference Theory", in James E. Alatis (ed.): *Contrastive Linguistics and Its Pedagogical Implications. Report of the Nineteenth Annual Round Table Meeting on Linguistics and Language Studies*. Washington D.C.: Georgetown University Press, pp. 113-122.
- Li, Ping and Yasuhiro Shirai (2000): *The Acquisition of Lexical and Grammatical Aspect*. Berlin and New York: Mouton de Gruyter.  
<http://dx.doi.org/10.1515/9783110800715>
- Montrul, Silvina and Roumyana Slabakova (2000): "Acquiring Semantic Properties of Preterit and Imperfect Tenses in L2 Spanish", in S. Catherine Howell, Sarah A. Fish, and Thea Keith-Lucas (eds.), *BUCLD 24 Proceedings*, 2. Somerville, Mass.: Cascadilla Press, pp. 534-545.

- Slabakova, Roumyana and Silvina Montrul (2002): "On Aspectual Shifts in L2 Spanish", in Barbora Skarabela, Sarah A. Fish, and Anna H.-J. Do (eds.), *BUCLD 26 Proceedings*, 2. Somerville, Mass.: Cascadilla Press, pp. 631-642.
- Odlin, Terrence (1989): *Language Transfer*. Cambridge: Cambridge University Press.  
<http://dx.doi.org/10.1017/CBO9781139524537>
- Ringbom, Hakan and Rolf Palmberg (1976): *Errors Made by Finns and Swedish-Speaking Finns in the Learning of English*. Abo, Finland: Department of English, Abo Akademi.
- Salaberry, Rafael (2000): *The Development of Past Tense Morphology in L2 Spanish*. Amsterdam and Philadelphia: John Benjamins.
- Schachter, Jacquelyn and William Rutherford (1979): "Discourse Function and Language Transfer", *Working Papers in Bilingualism* 19, pp. 1-12.
- Solé, Yolanda R. and Caros A. Solé (1977): *Modern Spanish Syntax: A Study in Contrast*. Lexington, Mass: Heath.
- Weinreich, Uriel (1974): *Language in Contact*. The Hague: Mouton.



## **The Effect of Acoustic Training on the Acquisition of Second Language Spanish Stops**

| John J. Stevens

| *University of North Carolina Wilmington*

**Abstract:** This investigation addresses the question of whether native English-speaking adult learners of Spanish as a second language (L2), who are trained to measure the voice onset time (VOT) of stop phonemes using speech analysis computer software (waveforms and spectrograms), show significantly more improvement in the pronunciation of the Spanish stops--/p t k b d g/--in terms of reduced VOT than do students who receive no instruction in acoustic analysis. An experiment was conducted that compared the production of the Spanish stops in a sentence context of an experimental group of learners who received acoustic training with that of a control group. Acoustic and quantitative analyses of speech data extracted from both a pretest and a posttest were performed, and the mean VOT values of the Spanish stops of both groups of learners were compared. The results revealed no significant differences in the reduction of VOT of the voiceless stops /p t k/ between the groups. However, a positive correlation between reduction in VOT and acoustic training was found for the voiced stop /b/, although not for /d/ and /g/.

**Keywords:** L2 Spanish acquisition; phonetics/phonology, language instruction, acoustic training.



## 1. INTRODUCTION

This paper addresses the question of whether instruction can lead to improvement in the pronunciation of a second language (L2). Specifically, it tests the hypothesis that a particular type of pedagogical intervention—training in basic acoustic analysis techniques—will lead to gains in the pronunciation of the Spanish stop consonants.

The acquisition of L2 phonology has generally not received the same amount of attention as other areas of second language acquisition (SLA) such as vocabulary, morphology, and syntax. Nevertheless, the literature does reveal an important body of research that has emerged over the last twenty five years, which underscores the continuing interest in this field (e.g., Archibald 1998; Best and Tyler 2007; Bongaerts, Plankton and Schils 1995; Bundgaard-Nielsen, Best and Tyler 2011; Flege 1991, 2009; Flege and Eefting 1988; Flege, Frieda, Walley and Randazza 1998; Larson-Hall 2006; Moyer 1999; Munro and Bohn 2007; Munro and Derwing 1999; Piske, MacKay and Flege 2001). Although this research examines a range of issues, including the acquisition of individual segments, prosodic features, phonetic characteristics of second languages, and factors that may intervene in the development of L2 phonology such as age, gender, motivation, and aptitude, relatively few studies address the issue of instruction itself. With the exception of Suter 1976, the empirical studies that have investigated the impact of phonological instruction generally have reported finding beneficial effects on L2 pronunciation (e.g., Bongaerts 1999; Couper 2003; Neufeld and Schneiderman 1980; Moyer 1999; Derwing, Munro and Wiebe 1997, 1998).

The handful of studies that have concentrated on L2 Spanish have also shown a positive correlation between instruction and improvement in pronunciation. Elliott (1997), for example, targeted a subset of Spanish sounds thought to be particularly difficult for American English speakers. He found overall significant improvement for an experimental group that received pronunciation instruction compared to a control group that received no such instruction. González-Bueno (1997), looking at the Spanish voiceless unaspirated stops /p t k/, found that an experimental group of intermediate-level Spanish language students that received regular phonetics instruction and pronunciation practice throughout the semester showed significant gains in reducing aspiration in the pronunciation of /p/ and /t/ compared to a control group. However, no improvement was reported for /k/. In an experiment that explored the relationship between the L2 production and

perception of the Spanish stop consonants /p/ and /b/, Zampini (1998) did not find a strong correlation between learners' perceptual capabilities and their ability to produce these sounds. Nevertheless, she did report that, after receiving explicit instruction, learners made significant gains in the production of /p/ but not /b/, suggesting that voiced stops may take longer to acquire than voiceless stops. Finally, Lord (2005) reported that students in an upper-division university Spanish phonetics class, who received standard phonetics instruction, practiced with voice analysis software and completed a self-analysis project designed to increase their awareness of differences between their own pronunciation and native pronunciation, made significant gains in the pronunciation of the trill /r/, diphthongs within and between words, and the fricative consonants [β ð γ]. Students also showed improvement in the pronunciation of /p t k/ in terms of reduced aspiration; however, this improvement proved not to be significant.

One of the central pedagogical issues to be debated in SLA over the last several years has been whether learners acquire an L2 better while engaged in meaning-focused activities or by form-focused instruction. Communicative language teaching (CLT) refers to a variety of approaches whose key ingredient is meaningful comprehensible input, which is provided by activities that emphasize the appropriate use of language rather than explicit knowledge of language. One type of CLT known as content-based instruction (CBI) has become particularly widespread and uses the target language to teach subject matter: Students learn the new language by using it within a specific academic context. Form-focused instruction, on the other hand, refers to teaching that draws the learner's attention to language form, either implicitly or explicitly, and includes techniques associated with traditional approaches to L2 instruction, such as the teaching of grammatical rules and corrective feedback (Spada 1997). Research has shown that while CLT appears to be effective for the development of overall fluency, comprehension skills, and communicative abilities, learners continue to have difficulties with pronunciation and morphological, syntactic, and pragmatic features (see, for example, Harley and Swain 1984; Lyster 1987). Studies that have examined the inclusion of form-focused instruction within CLT and CBI classrooms have led to a growing consensus that a combination of form and meaning is more effective than the exclusive use of either one (Spada and Lightbown 2008).

Some researchers have endeavored to tease out and test the effect of explicit information in conjunction with form-focused instruction on L2 morphosyntactic development (e.g., Norris and Ortega 2000; Spada and

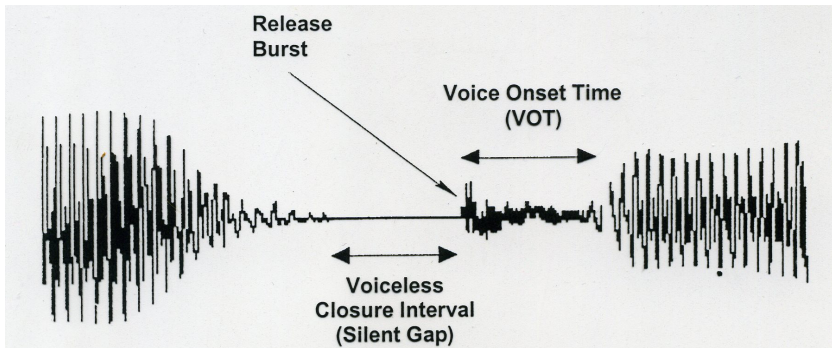
Tomita 2010). Explicit information is thought to enhance L2 acquisition by making learners aware of target language features that are particularly problematic “either because they are acoustically imperceptible (e.g., most grammatical morphology in English) or redundant and unlikely to affect comprehension (e.g., word order in English questions)” (Spada and Lightbown 2008, p. 186). In the context of L2 phonological development, Derwing and Munro (2005) emphasized the importance of explicit information in the teaching of pronunciation claiming that “students learning L2 pronunciation benefit from being explicitly taught phonological form to help them notice the difference between their own productions and those of proficient speakers in the L2 community” (p. 388). More recently, in a study that investigated the acquisition of the English liquid /ɹ/ by Japanese learners, Saito (2013) examined the effect of adding explicit phonetic information to meaning-oriented lessons that focused on form. Hypothesizing that explicit information would promote learners’ awareness of the perceptual difference between a new L2 sound and its native, or first language (L1), counterpart, Saito found that learners who received form-focused instruction enhanced with explicit information outperformed learners who received only form-focused instruction in their ability to pronounce /ɹ/.

The present investigation seeks to contribute to the above research by examining the impact on L2 Spanish pronunciation of the addition of acoustic training to phonetic instruction that is form-focused. It is predicted that because acoustic training will provide learners with explicit information that will allow them to notice differences between the L1 and L2 phonological systems that are difficult for them to perceive, they will show greater improvement in L2 pronunciation than learners who receive no such explicit phonetic training. In order to test this hypothesis, a group of learners that received both pronunciation instruction and acoustic training will be compared to a similar group of learners that received pronunciation instruction only in the articulation of a subset of sounds considered to be particularly difficult for English-speaking learners, the Spanish stop consonants.

Both English and Spanish have a series of stop phonemes that are distinguished by voicing: voiceless /p t k/ contrasted with voiced /b d g/. However, significant differences exist between the two languages regarding *voice onset time* (VOT). VOT refers to the duration of time, measured in milliseconds (ms), which occurs between the release burst of the stop and the beginning of vocal chord vibration (see Figure 1). Whereas English /p t k/ are said to be “long-lag” stops realized with relatively long VOT values,



especially word-initially and in the onset of a stressed syllable, /b d g/ are “short-lag” stops produced with considerably shorter VOT values. In contrast, the Spanish voiceless stops are produced with short VOT values; however, their voiced counterparts are pronounced with prevoicing (or lead voicing) in which the vocal folds begin vibrating before the release of the articulators.



**Figure 1.** Waveform of a Syllable-Initial Stop Consonant in English, Showing Voiceless Closure Interval, Release Burst, and Voice Onset Time (VOT) (adapted from Zampini and Green 2001).

Because the release burst (the moment when the articulators separate) is marked as 0.00 on the time scale, prevoicing is represented by a negative number, which indicates the amount of time the vocal folds have been vibrating prior to the release of the stop. These contrasts are represented in Figures 2A and 2B of Figure 2, which show sample waveforms and VOT values for the stops /p/ and /b/ in the words *pole* and *bowl*, respectively, spoken by a native speaker of English. Sample waveforms and VOT values for /p/ and /b/ in the words *polo* and *bola*, produced by a native speaker of Spanish, appear in Figures 2C and 2D, along with their respective VOT values.

Table 1 further illustrates the differences between the English and Spanish stops, providing the mean VOT production values as well as the range of VOT values for English and Spanish stops as reported in a seminal study by Lisker and Abramson (1964). Despite sharing the same phonemic representation, /p t k b d g/, Table 1 reveals important differences between the English and Spanish stops in terms of VOT. The relatively long-lag VOTs with which the English voiceless stops are produced cause aspiration; however, the Spanish voiceless stops are never aspirated and share short-lag VOT values similar to the English voiced stops. This is why inexperienced English speakers sometimes perceive Spanish /p/ as /b/; for example, mistaking *peso* (‘weight’) for *beso* (‘kiss’). Finally, while the Spanish voiced

stops are prevoiced, their English equivalents are usually pronounced with short-lag VOT values, although they may be sometimes prevoiced.

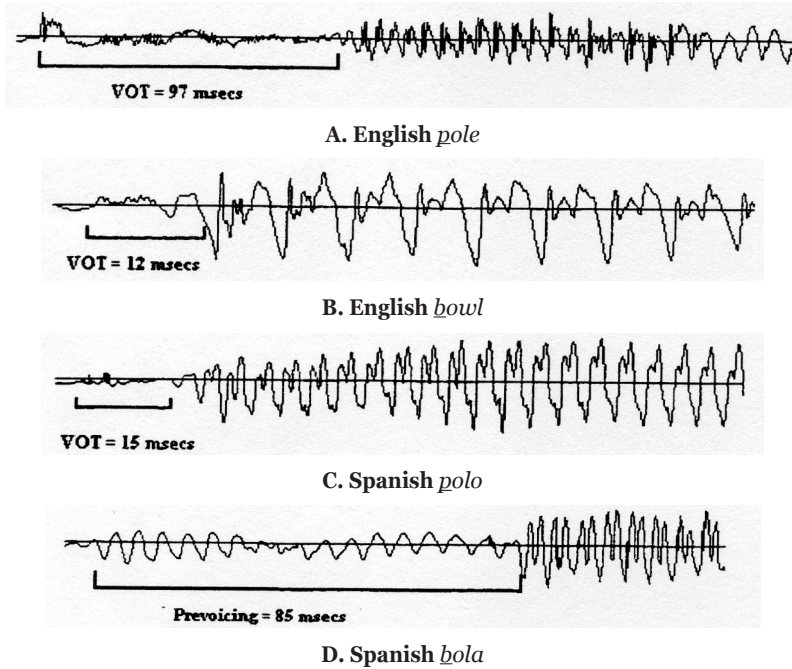


Figure 2. Waveforms for English /p/ and /b/ and Spanish /p/ and /b/.

English			Spanish		
Stop	Mean VOT	Range	Stop	Mean VOT	Range
/p/	58	20 / 120	/p/	4	0 / 15
/t/	70	30 / 105	/t/	9	0 / 15
/k/	80	50 / 135	/k/	29	15 / 55
/b/	1	0 / 5	/b/	-138	-235 / -60
/d/	5	0 / 25	/d/	-110	-170 / -75
/g/	21	0 / 35	/g/	-108	-165 / -45

Table 1. Mean VOT Values in (in ms) for English and Spanish Stops (Lisker and Abramson 1964)

Note. Positive values indicate voicing lag; negative values indicate prevoicing.

The specific hypothesis to be tested in the present study is the following: American English-speaking adult learners of Spanish who receive acoustic training in addition to pronunciation instruction will exhibit significant improvement in terms of the reduction in VOT of the Spanish stops compared to learners who are provided with pronunciation instruction only, because the acoustically trained learners will have a greater opportunity to notice and become more aware of the differences between the English and Spanish stops, which in turn will lead to the formation of new L2 phonetic categories and thus enhance their ability to produce these sounds.

## 2. METHOD

### 2.1. Participants

In order to test the above-stated hypothesis, an experimental study was designed to examine the production of word-initial Spanish stop consonants by American university students enrolled in an upper-division Spanish phonetics course at the University of North Carolina Wilmington. After eliminating potential participants from the original subject pool, either because they were non-native speakers of English, had studied Spanish abroad, had been absent for the experimental sessions, or else exhibited a level of proficiency in Spanish deemed unsuitable for the purposes of this experiment, two subject groups were formed comprising five students each: an experimental group that received pronunciation instruction and acoustic training and a control group that received pronunciation instruction only. All of the participants were female native speakers of English, ranging in age from 19 to 23 years. None reported knowing, or having studied, another language besides Spanish, and all denied having received any specific instruction in Spanish pronunciation. The amount of previous formal instruction in Spanish was similar for the two groups, averaging 3.5 years for the control group and 3 years for the experimental group. Interviews based on ACTFL's Oral Proficiency Interview (OPI) revealed that the subjects shared a similar level of proficiency in Spanish, which ranged from intermediate low to intermediate high according to ACTFL criteria<sup>1</sup>.

---

<sup>1</sup> These interviews were conducted by the author, who is trained in administering and scoring the OPI.

## 2.2. Instrument and Procedure

Sample recordings of the subjects' pronunciation of the Spanish stops while reading a randomized list of sentences were taken from both a pretest and a posttest. The elicitation instrument consisted of 60 randomized target words containing 10 occurrences each of the six Spanish stops in stressed syllables in absolute word-initial position. To control for the possible effect of the following vocalic segment, each stop occurred twice in different words containing each of the Spanish vowels--/i u e o a/. This rendered a total of 60 stop tokens per subject for each test. The pretest recording occurred during the second week of the course. At this point, the learners had only begun to study basic phonetic concepts, as well as Spanish syllable structure and orthographic rules. The posttest recording was obtained later in the semester, three weeks after both the control and experimental groups had received pronunciation instruction in the Spanish stop consonants and approximately two weeks after the experimental group had received explicit instruction in the form of acoustic training. The control and experimental groups were enrolled in different sections of the same phonetics course. Both sections were taught by the same instructor (the author), who took great care in ensuring that the pronunciation instruction was identical for both groups. This instruction was based on the information contained in the course textbook *iQué bien suena!* (Stokes 2005), which describes the stop phones in terms of place and manner of articulation, emphasizing the difference between the articulation of these sounds in English and Spanish. This textbook does not describe the differences between the English and Spanish stops in terms of long-lag vs. short-lag categories and VOT. It does, however, emphasize that L2 Spanish learners should refrain from aspirating /p t k/. This instruction was supplemented by perception exercises in which learners were asked to discriminate between minimal pairs of Spanish words pronounced by a native speaker that differed only in the voicing of the syllable-initial stop; for example, *basta/pasta*, *domar/tomar*, *guiso/quiso*. These perception exercises were followed by a series of production exercises. After pronouncing the Spanish stops together as a class, first in individual words and then in sentences, learners moved on to pair work, where they took turns as speaker and listener. For these exercises, the instructor focused learners' attention on suppressing the aspiration of /p t k/ by decreasing the amount of air in the production of these sounds. In regard to the pronunciation of /b d g/, learners were instructed to begin the vibration of their vocal chords before the separation of the articulators.

In addition to this form-focused instruction, the experimental group subsequently received explicit phonetic information in the form of acoustic training. First the learners received information about the acoustic characteristics of the English and Spanish stops through the aid of computer-projected images in the classroom. Using waveforms and spectrograms, the instructor explained the difference between long-lag, short-lag, and prevoiced stops, paying particular attention to the differences in duration of VOT. The learners then spent two hour-long sessions in the university's Foreign Language Resource Center (language lab), where they received hands-on instruction in generating waveforms and spectrograms and in measuring VOT using the *Speech Analyzer* computer software program. After the initial training sessions, the subjects completed exercises in which they recorded themselves pronouncing words in English and Spanish that contained initial stops. The learners measured the VOT of these stops, and were then asked to write a paragraph comparing the differences in their English and Spanish stop production. Finally, the experimental subjects recorded themselves reading a list of 24 Spanish words containing 4 occurrences each of the six Spanish stops in initial position. They then measured the VOT of these stops and calculated the average VOT for each of the six stops.

### **2.3. Acoustic and Quantitative Analysis**

The speech data obtained from the pretest and posttest were digitized and analyzed acoustically using *Speech Analyzer*. Waveforms and spectrograms were generated and inspected for each stop. The VOT value of each token was measured beginning with the release of the articulators and ending with the onset of vocal chord vibration. In the case of prevoiced stops, the duration of lead voicing was measured as a negative VOT value. The reliability of the VOT measurement was assessed by creating a second set of waveforms and spectrograms for each of the six stops produced by five randomly selected subjects. The VOT values of these 30 stops were measured and compared to the values of the original waveforms/spectrograms. The mean difference between the two sets of measurements was minimal, averaging 3.5 ms, with a range of 0.00 to 5.0 ms.

Statistical analyses of the data were conducted using the *PASW 18* statistics program. Independent-samples *t* tests were run to compare the Spanish stop VOT values of the experimental and control subjects prior to the pedagogical intervention and to determine whether the amount of change in VOT was significantly different between the groups on the posttest. Paired-

samples *t* tests were conducted to measure whether VOT values differed significantly among the groups between the pretest and the posttest.

### 3. RESULTS

In order to determine the comparability of the control and experimental groups' Spanish stop production, independent samples *t* tests were performed for the six Spanish stops prior to the pedagogical intervention. The results are shown in Table 2, which gives the mean VOT values of each stop and indicates whether these values were statistically different between the groups at the outset of the experiment. These results reveal that there were no significant differences found between the groups for any of the mean VOT values of the respective Spanish stops, with the exception of the voiced stops /b/ and /d/, whose VOT values were significantly longer for the experimental group. Despite these differences, the overall similarity of the control and experimental groups' Spanish stop pronunciation renders them comparable for the purposes of the present investigation.

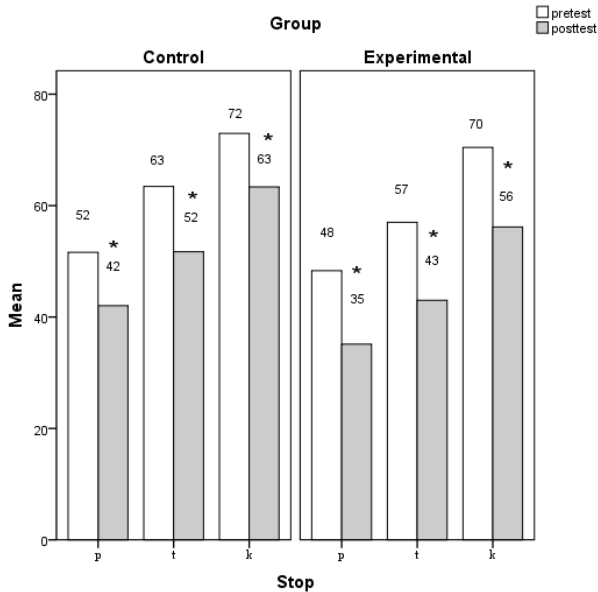
	Control	Experimental	
Stop	Mean VOT	Mean VOT	Difference
/p/	51.60	48.32	NS
/t/	63.42	57.00	NS
/k/	72.48	70.54	NS
/b/	14.66	19.06	*
/d/	19.96	25.54	*
/g/	39.18	36.38	NS

**Table 2.** Mean Pretest VOT Values (in ms) for Spanish Stops, Control and Experimental Groups (Note. NS indicates non-significant differences between pairs of means; an asterisk [\*] indicates significance<sup>2</sup>).

A series of paired-samples *t* tests was conducted for each of the Spanish stops in order to evaluate whether the subjects had improved their pronunciation of these sounds in terms of a reduction in VOT between the pretest and the posttest. The results of the paired-samples *t* tests for

<sup>2</sup> An alpha level of .05 was used for all statistical tests. Per convention, p-values of .000 are reported as "p < .001".

the voiceless stops /p t k/, represented graphically in Figure 3, show that both the control and the experimental groups experienced a significant mean reduction in VOT for all three stops<sup>3</sup>. For the control group, the mean difference for /p/ was -9.56 ms ( $t = 2.96, p = .005$ ); for /t/, the mean difference was -11.71 ms ( $t = 4.32, p < .001$ ); and for /k/, the mean difference was -8.96 ms ( $t = 3.58, p = .001$ ). For the experimental group, the mean difference for /p/ was -13.20 ms between the pretest and the posttest ( $t = 4.72, p < .001$ ); for /t/, the mean difference was -14.00 ms ( $t = 4.15, p < .001$ ); and for /k/, the mean difference was -14.27 ms ( $t = 5.71, p < .001$ ). Although the experimental group exhibited a larger reduction in VOT for each of the respective stops than did the control group (for /p/, the mean difference was 3.64 ms; for /t/, the mean difference was 2.29 ms; and for /k/, the mean difference was 4.68 ms), separate independent samples *t* tests revealed that the amount of reduction between the groups was nonsignificant for each of the three stops (for /p/,  $t = .85, p = .396$ ; for /t/,  $t = .53, p = .600$ ; and for /k/,  $t = 1.33, p = .186$ ).

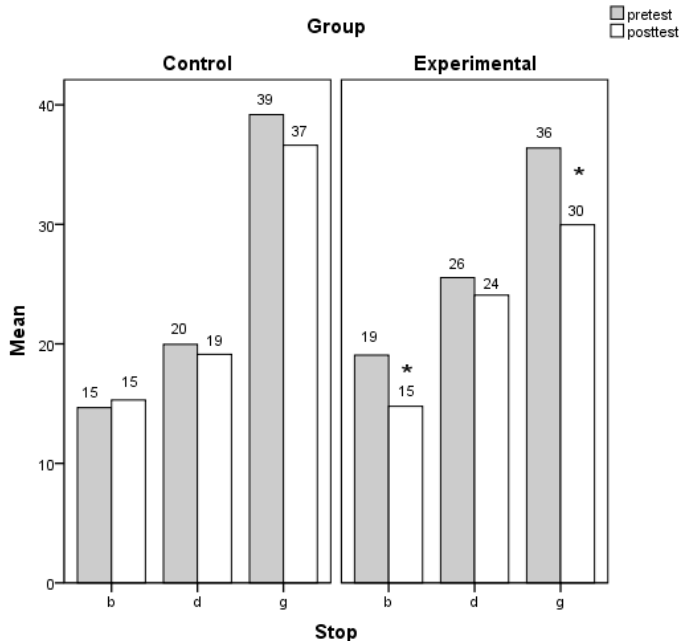


**Figure 3.** Mean Reduction in VOT (in ms) for /p t k/, by Group  
(Note. An asterisk [\*] indicates significance).

<sup>3</sup> For ease of exposition, the mean values displayed in all figures have been rounded to the nearest whole integer. For example, 51.60 has been rounded up to 52; 42.04 has been rounded down to 42.



The results of the paired-samples  $t$  tests for the voiced stops /b d g/ are represented graphically in Figure 4. These results show that the control group experienced a minimal reduction in VOT for /d/ of  $-.84$  ms and a slight reduction for /g/ of  $-2.56$  ms between the pretest and the posttest. None of these decreases was significant (for /d/,  $t = .77$ ,  $p = .443$ ; for /g/,  $t = 1.74$ ,  $p = .088$ ). On the other hand, the control group registered a minimal nonsignificant increase in VOT for /b/ of  $.64$  ms ( $t = .76$ ,  $p = .449$ ). At the same time, the experimental group evidenced a significant reduction in VOT for /b/ and /g/ of  $-4.29$  ms and  $-6.42$  ms, respectively (for /b/,  $t = 5.16$ ;  $p < .001$ ; for /g/,  $t = 3.01$ ,  $p = .004$ ). Although /d/ also experienced a decrease in VOT of  $-1.46$  ms, this reduction was nonsignificant ( $t = 1.84$ ;  $p = .071$ ). Independent samples  $t$  tests indicated that whereas the mean difference in VOT for /b/ between the pretest and the posttest was significant between the control and experimental groups (the mean difference was  $4.93$  ms,  $t = 4.18$ ,  $p < .001$ ), the difference in the amount of reduction in VOT for /d/ and /g/ between the groups was nonsignificant (for /d/, the mean difference was  $.62$  ms,  $t = .46$ ,  $p = .646$ ; for /g/, the mean difference was  $3.86$  ms,  $t = 1.49$ ,  $p = .139$ ).



**Figure 4.** Mean Reduction in VOT (in ms) for /b d g/, by Group  
(Note. An asterisk [\*] indicates significance).



Table 3 summarizes the results, giving the average VOT values as well as the minimum and maximum VOT values for each of the six Spanish stops for the control and the experimental groups on the pretest and the posttest. These results reveal that both groups experienced a significant reduction in VOT for each of the Spanish voiceless stops /p t k/. Although the experimental group achieved mean VOT values that were lower than those of the control group, the difference in the amount of VOT reduction between the groups was nonsignificant for each of the voiceless stops. In regard to the voiced stops, the experimental group showed a significant reduction in VOT for /b/ and /g/ and a slight nonsignificant decrease for /d/. For its part, the control group experienced a small nonsignificant reduction in VOT for /d/ and /g/, but showed a nonsignificant increase for /b/.

Control					Experimental				
Pretest			Posttest		Pretest			Posttest	
Stop	VOT	Range	VOT	Range	Stop	VOT	Range	VOT	Range
/p/	51.60	14/97	42.04	10/94	/p/	48.32	17/114	35.12	7/99
/t/	63.43	25/99	51.71	17/78	/t/	57.00	31/137	43.00	9/108
/k/	72.34	43/105	63.38	36/102	/k/	70.44	45/123	56.17	23/89
/b/	14.66	-5/27	15.30	9/28	/b/	19.06	7/32	14.77	-5/26
/d/	19.96	11/39	19.12	11/36	/d/	25.54	7/38	24.08	-4/39
/g/	39.18	23/77	36.62	17/55	/g/	36.38	21/92	29.96	-16/48

**Table 3.** Mean VOT Values (in ms), by Group.

#### 4. DISCUSSION

The results of the quantitative analysis only partially support this study's original hypothesis: Whereas both the control and experimental groups significantly improved their pronunciation of the voiceless stops /p t k/ in terms of a reduction in VOT, only the experimental group showed a significant improvement in the voiced stops /b/ and /g/. Furthermore,

the only stop for which the experimental group showed significantly more improvement than the control group was /b/; however, the VOT values for this stop were not significantly different between the groups on the posttest.

It seems clear that the form-focused instruction both groups of learners received resulted in improvement of their L2 Spanish stop production. This finding not only supports the general consensus that instruction can lead to improvement in L2 pronunciation, but also confirms the findings of previous studies that showed a positive correlation between instruction and improvement in the pronunciation of the L2 Spanish stops (see, for example, González-Bueno 1997; Lord 2005; Zampini 1998). In this case, the type of instruction the learners received—information on manner and place of articulation, and perception and production exercises—was sufficient for them to make gains in their pronunciation of /p t k/ by reducing aspiration, although neither group achieved VOT values approximating native speaker norms as reported in Table 1<sup>4</sup>.

Flege's Speech Learning Model (1995, 2003, 2009) relates phonological production to perception by claiming that adult learners perceive L2 sounds in terms of "new" and "similar" categories. "New" L2 sounds have no L1 counterpart and are easily distinguished from L1 sounds. Learners have less difficulty producing new L2 sounds accurately because they are able to create new phonetic categories for them. On the other hand, "similar" L2 sounds differ systematically from an easily identifiable L1 counterpart, yet are perceived by learners as belonging to the same native language phonetic category. Learners experience more difficulty producing similar L2 sounds because they fail to establish new phonetic representations for them that are distinct from those found in the L1. Although Flege does not claim that all L2 production errors are perceptually motivated, he does emphasize the importance of phonetic-level noticing of perceptual aspects of new segmental sounds in order for learners to create new phonetic categories and consequently increase their ability to produce new L2 sounds.

In terms of Flege's model, both the Spanish voiceless and voiced stops represent "similar" L2 sounds: Although they differ systematically from their English counterparts, learners perceive them as belonging to

---

<sup>4</sup> There can be considerable variation in VOT values according to the recording equipment and acoustic analysis software used. Reeder (1998), for example, reported mean native speaker VOT values of 17 ms for /p/, 17 ms for /t/, and 28 ms for /k/, while Lord (2005) reported values of 21 ms for /p/, 22 ms for /t/, and 33 ms for /k/. Be that as it may, the learners in the present study did not come close to any of the reported native target VOT values.

the same native language phonetic category. Flege's model predicts that learners will have trouble producing these sounds accurately because they have difficulty creating new phonetic categories for them. However, in the case of the voiceless stops /p t k/, learners have the advantage that each of these stops has an unaspirated variant that occurs in English obligatorily after an /s/ in the same syllable (e.g., *stone* [stóʊn]), obligatorily in syllable-final position (e.g., *artful* [áɪt.fəl]), and optionally as the last sound in a breath group where it alternates with an unreleased stop allophone (e.g., *cat*, [k<sup>h</sup>æt] ~ [k<sup>h</sup>æt̚]). Thus, the learners already possessed in their English L1 phonological inventory the L2 target sounds to be acquired. Here, the form-focused instruction that both groups of learners received appears to have been sufficient for them to make significant progress in restructuring their interlanguage phonological system such that they began to produce unaspirated stops in word-initial stressed syllables, an environment where these sounds are aspirated in English.

Compared to the reduction in aspiration of the voiceless stops, the prevoicing associated with the Spanish voiced stops /b d g/ was more challenging for the learners to achieve; for example, the control group did not produce any stops with prevoicing on the posttest, while the experimental group only produced one token each of prevoiced /b d g/ on the posttest. Although the control and experimental groups both experienced a slight improvement in the pronunciation of /d/, only the experimental group made significant progress in reducing the VOT of /b/ and /g/; however, these VOT values did not come close to the (prevoiced) target values as reported in Table 1.

Even though the voiced stops /b d g/ are considered to be similar L2 sounds according to Flege's model, the acquisition of their prevoiced allophones seems to be more difficult than the acquisition of the unaspirated voiceless stops: Whereas for [p t k] learners need simply extend the use of L1 English allophones to new L2 Spanish environments, for [b d g], they must acquire a new articulation for these sounds, that is, prevoicing. Furthermore, compared to detecting the presence or absence of aspiration in voiceless stops, learners find it more difficult perceiving the prevoicing of Spanish voiced stops. The experimental learners in this investigation, who received form-focused pronunciation instruction in conjunction with explicit information in the form of acoustic training, did make some progress in reducing the VOT of the voiced stops; however, this improvement was significantly better than that of the control group only in the case of /b/.

In the typical university Spanish phonetics classroom, form-focused instruction and explicit information are not mutually exclusive: Students are explicitly taught about the Spanish sound system, and their attention is focused on phonological form in order for them to notice the differences between their own pronunciation and those of native speakers of Spanish. In this study, the control learners received form-focused instruction that included a certain amount of explicit information: Learners received information on place and manner of articulation and were taught that they should extend the use of the unaspirated variants of /p t k/, which they already pronounced in English after /s/ and as the last sound in a breath group, to all the environments where these stops occur in Spanish, especially word-initially and before a stressed vowel. The experimental group, on the other hand, received acoustic training, which entailed not only more form-focused instruction, but also more detailed, explicit information about the phonetic and phonological differences between the English and Spanish stops. This information provided the context for the experimental learners to analyze their own English and Spanish stop production by measuring VOT in waveforms and spectrograms. This additional focus is believed to have led to learners' increased awareness of the difference between the English short-lag voiced stops and the Spanish prevoiced stops which, in turn, may have been responsible for their significant reduction in VOT of /b/ compared to the pronunciation only learners.

## 5. CONCLUSION

This study has shown that English-speaking learners who received acoustic training in addition to traditional form-focused pronunciation instruction experienced more improvement in the articulation of Spanish /b/ in terms of reduced VOT than did learners who received pronunciation instruction only. However, only one of the six sounds examined here improved, and the VOT values for /b/ of the experimental learners were not significantly lower than those of the control group on the posttest. Thus, it appears that acoustic training, in this case, was of little to no benefit.

Although this investigation found that acoustic training did not provide sufficient improvement in the Spanish stops to warrant its inclusion in the phonetics classroom, more research is needed to confirm this finding and to address this study's limitations. For example, future research could

employ larger subject groups consisting of both females and males in order to obtain more conclusive statistical results. Also, phonetic samples elicited in a more informal conversational style would render pronunciation data more typical of learners' everyday L2 speech. Finally, future investigations could examine the effect of acoustic training on the acquisition of other features of L2 Spanish that tend to pose problems for English-speaking learners, such as the stop/fricative alternation of /b d g/, vowel length, and the vibrant trill /r/. Such research would serve to inform an instructor's decision as to whether acoustic training leads to sufficient improvement in L2 pronunciation to justify the inclusion of such training in introductory Spanish phonetics courses.

## REFERENCES

- Archibald, John (1998): "Second language phonology, phonetics, and typology", *Studies in Second Language Acquisition* 20.1, pp. 189-211.
- Best, Catherine and Michael Tyler (2007): "Nonnative and second-language speech perception", in Ocke-Schwen Bohn and Murray J. Munro (eds.): *Language Experience in Second Language Speech Learning: In Honour of James Emil Flege*. Amsterdam: Benjamins, pp. 13-34.  
<http://dx.doi.org/10.1075/llt.17.07bes>
- Bongeaerts, Theo (1999): "Ultimate attainment in L2 pronunciation: The case of very advanced late learners", in David Birdsong (ed.): *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Erlbaum, pp. 133-160.
- and Brigitte Plankton, and Erik Schils (1995): "Can late learners attain a native accent in a foreign language? A test of the Critical Period Hypothesis", in David M. Singleton and Zsolt Lengyel (eds.): *The Age Factor in Second Language Acquisition*. Clevedon: Multilingual Matters, pp. 30-50.
- Bundgaard-Nielsen, Rikke, Catherine Best and Michael Tyler (2011): "Vocabulary size is associated with second-language vowel perception performance in adult learners", *Studies in Second Language Acquisition* 33, pp. 433-461.  
<http://dx.doi.org/10.1017/S0272263111000040>
- Couper, Graeme (2003): "The value of an explicit pronunciation syllabus in ESOL teaching". *Prospect* 18.3, pp. 53-70.
- Derwing, Tracey, and Murray Munro (2005): "Second language accent and pronunciation teaching: A research-based approach". *TESOL Quarterly* 39, pp. 379-397.  
<http://dx.doi.org/10.2307/3588486>

- Derwing, Tracey, Murray Munro and Grace Wiebe (1997): "Pronunciation instruction for fossilized learners: Can it help?" *Applied Language Learning* 8, pp. 217-235.
- Murray Munro and Grace Wiebe (1998): "Evidence in favor of a broad framework for pronunciation instruction". *Language Learning* 48, pp. 393-410.  
<http://dx.doi.org/10.1111/0023-8333.00047>
- Elliott, Raymond (1997): "On the teaching and acquisition of pronunciation within a communicative approach". *Hispania* 80, pp. 95-108.  
<http://dx.doi.org/10.2307/345983>
- Flege, James (1991): "Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language". *Journal of the Acoustical Society of America* 89.1, pp. 395-411.  
<http://dx.doi.org/10.1121/1.400473>
- (1995): "Second language speech learning theory: Theory, findings, and problems", in Winifred Strange (ed.): *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York, pp. 233-272.
- (2003): "Assessing constraints on second-language segmental production and perception", in Antje Meyer and Niels Schiller (eds.): *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*. Berlin: Mouton de Gruyter, pp. 319-355.
- (2009): "Give input a Chance!", in Thorsten Piske and Martha Young-Scholten (eds.): *Input Matters in SLA*. Bristol, UK: Multilingual Matters, pp. 175-190.
- and Wieke Eefting (1988): "Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation". *Journal of the Acoustical Society of America* 83.2, pp. 729-740.  
<http://dx.doi.org/10.1121/1.396115>
- Elaina Frieda, Amanda Walley and Lauren Randazza (1998): "Lexical factors and segmental accuracy in second language speech production". *Studies in Second Language Acquisition* 20.1 (1998): 155-187.

- González-Bueno, Manuela (1997): "The effects of formal instruction on the acquisition of Spanish stop consonants", in William Glass and Ana Teresa Pérez-Leroux (eds.): *Contemporary Perspectives on the Acquisition of Spanish*. Somerville, MA: Cascadilla, pp. 57-75.
- Larson-Hall, Jennifer (2006): "What does more time buy you? Another look at the effects of long-term residence on production accuracy of English /r/ and /l/ by Japanese speakers". *Language and Speech* 49, pp. 521-548.  
<http://dx.doi.org/10.1177/00238309060490040401>
- Lisker, Leigh, and Arthur Abramson (1964): "A cross-language study of voicing in initial stops". *Word* 20, pp. 384-422.
- Lord, Gillian (2005): "(How) can we teach foreign language pronunciation? On the effects of a Spanish phonetics course". *Hispania* 88, pp. 557-567.  
<http://dx.doi.org/10.2307/20063159>
- Moyer, Arlene (1999): "Ultimate attainment in L2 phonology: The critical factors of age, motivation and instruction". *Studies in Second Language Acquisition* 21.1, pp. 81-108.  
<http://dx.doi.org/10.1017/S0272263199001035>
- Munro, Murray and Ocke-Schwen Bohn: (2007) (eds.): *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. Amsterdam & Philadelphia: Benjamins.
- and Tracey Derwing (1999): "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners". *Language Learning* 49, pp. 285-310.  
<http://dx.doi.org/10.1111/0023-8333.49.s1.8>
- Neufeld, Gerald and Eta Schneiderman (1980): "Prosodic and articulatory features in adult language learning". *Canadian Modern Language Review* 34, pp. 105-109.
- Norris, John and Lourdes Ortega (2000): "Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis". *Language Learning* 50, pp. 417-528.  
<http://dx.doi.org/10.1111/0023-8333.00136>
- Piske, Thorsten, Ian MacKay and James Flege (2001): "Factors affecting degree of foreign accent in an L2: A review". *Journal of Phonetics* 29, pp. 191-215.  
<http://dx.doi.org/10.1006/jpho.2001.0134>



- Reeder, Jeffrey (1998): "English speakers' acquisition of voiceless stops and trill in L2 Spanish", in Dale Koike and Mark Carpenter (eds.): *Texas Papers in Foreign Language Education*. U. of Texas, Austin: Foreign Language Education Program, pp. 101-118.
- Saito, Kazuya (2013): "Reexamining effects of form-focused instruction on L2 pronunciation". *Studies in Second Language Acquisition* 35.1, pp.1-29.  
<http://dx.doi.org/10.1017/S0272263112000666>
- Spada, Nina (1997): "Form-focused instruction and second language acquisition: A review of classroom and laboratory research". *Language Teaching* 30, pp. 73-87.  
<http://dx.doi.org/10.1017/S0261444800012799>
- and Patsy Lightbown (2008): "Form-focused instruction: Isolated or integrated?" *TESOL Quarterly* 42.2, pp. 181-207.
  - and Yasuyo Tomita (2010): "Interactions between type of instruction and type of language feature: A meta-analysis". *Language Learning* 60, pp. 263-308.  
<http://dx.doi.org/10.1111/j.1467-9922.2010.00562.x>
- Stokes, Jeffrey (2005): *iQué bien suena! Mastering Spanish Phonetics and Phonology*. Boston: Houghton Mifflin.
- Suter, Richard (1976): "Predictors of pronunciation accuracy in second language learning". *Language Learning* 26, pp. 233-253.  
<http://dx.doi.org/10.1111/j.1467-1770.1976.tb00275.x>
- Zampini, Mary (1998): "The relationship between the production and perception of L2 Spanish stops", in Dale Koike and Mark Carpenter (eds.): *Texas Papers in Foreign Language Education*. U. of Texas, Austin: Foreign Language Education Program, pp. 85-100.
- and Kerry Green: (2001): "The voicing distinction in English and Spanish: The relationship between perception and production", in Janet Nicol (ed.): *One Mind, Two Languages: Bilingual Language Processing*. Malden, MA: Blackwell, pp. 23-48.

**IV**  
**Language in Society**





## **Textos fundacionales de la indignación: Aproximación desde el análisis del discurso**

| Francisco Moreno Fernández  
| *Universidad de Alcalá*

**Abstract:** The intent of this paper is to pay attention to some foundational speeches of the outraged movement, especially in Spain. The communicative dimension of outraged movement presents speeches and messages from those who are considered their ideologues. The methodology used for this preliminary approach to the “outraged language” is discourse analysis. It is an empirical analysis, where the speech is not conceived as a sequence of linguistic units, but as a set of linguistic forms and meanings that interact with social and cultural meanings.

**Keywords:** Indignated, revolution, 15M, discourse analysis, counterdiscourse, politics.



## 1. INTRODUCCIÓN

Se conoce como *movimiento de los indignados*, *spanish revolution* o *#15M* al movimiento social que expresa la indignación ciudadana contra un sistema de partidos políticos que no se considera representativo, un orden económico calificado de injusto y una organización social que merma las libertades individuales y penaliza a los desprotegidos. Estas páginas se refieren a un universo de ideas y acciones vinculadas al movimiento de los indignados, entre las que reclaman específicamente nuestra atención las relacionadas con la lengua y con la comunicación; las relativas a lo que podríamos llamar el discurso de la indignación.

Para Manuel Castells (2012), los movimientos de protesta que han sacudido el Occidente y los países árabes durante los últimos años se gestan en las redes informáticas y cuajan en los espacios públicos ocupados, normalmente enclaves urbanos de especial significación dentro de sus ciudades: la puerta del Sol de Madrid, la plaza Cataluña de Barcelona, Wall Street en Nueva York, la plaza Tahrir de El Cairo o las calles de Túnez, como gran escenario de la revolución del Jazmín. Pero tanto en los espacios físicos como en los ciberespacios, la comunicación y los discursos —en definitiva, la lengua— cumplen una función angular (Martín Rojo 2012). La intención de este breve ensayo es llamar la atención sobre algunos discursos fundacionales del movimiento indignado, sin pretender el análisis de otros aspectos lingüísticos o comunicativos tratados por otros autores (Martín Corvillo 2011; Pujante y Morales 2013).

La dimensión comunicativa del movimiento indignado tiene entre sus exponentes más destacados los discursos y mensajes de los que se consideran sus ideólogos. Se trata de intelectuales —literatos, ensayistas, profesores, profesionales liberales— que fueron sacando a la luz escritos o declaraciones que alimentaron de contenido los movimientos indignados y contribuyeron a su andadura. Como consecuencia parcial de ello, se generaron contradiscursos y contramensajes, con el propósito de desprestigiar o mitigar los efectos de los primeros, que se erigieron en argumentario de una corriente de opinión contraria a los intereses de los indignados y que se difundió principalmente a través de medios de comunicación social. Unos discursos y otros, los mensajes indignados y sus contramensajes, son el material de construcción —y, a la vez, la manifestación— de dos corrientes ideológicas de signo opuesto. En la línea de la retórica de la construcción social de Jonathan Potter (1998), unos y otros quedarían configurados

como repertorios argumentales a favor o en contra de unas realidades, unas ideas o unos procesos, pero al tiempo constituirían el medio desde el cual se posicionan los propios intereses y se desacreditan los contrarios, conforme se destacan las credenciales e intenciones propias, buscando finalmente un mayor prestigio y una mayor verosimilitud de cada posicionamiento ante el conjunto de la sociedad.

El método utilizado para esta preliminar aproximación al lenguaje de los indignados es el del análisis del discurso, tal y como fue formulado por Deborah Schiffrin (1994). Se trata de un análisis de naturaleza empírica, en el que el discurso no se concibe como una simple secuencia de unidades lingüísticas, sino como un conjunto de formas y significados lingüísticos que interactúan con significados sociales y culturales, de tal manera que las estructuras y los significados discursivos se despliegan interactivamente dentro un medio sociocultural. Al mismo tiempo, las expresiones que componen el discurso se producen e interpretan en situaciones concretas, interpretables desde las teorías del discurso como interacción social de Teun van Dijk (2008; 2011). Las elecciones de los hablantes se producen según sus intenciones, las estrategias convencionales, las formas y significados lingüísticos y los contextos sociales. Todo ello dentro de unos marcos culturales de creencias y percepciones (Moreno Fernández 2012).

## 2. TEXTOS FUNDACIONALES

Los discursos sociales, morales y regeneradores, la denuncia de los abusos de poder o la crítica a los mecanismos económicos y financieros más deshumanizados no son, ni mucho menos, patrimonio del movimiento indignado, ni tuvieron una relación directa con su origen. Los alegatos de Ernesto Sábato por la resistencia, sus discursos contra las dictaduras, la infinidad de declaraciones de José Saramago —«el poder lo contamina todo; es tóxico», decía con frecuencia— o las ideas de un sociólogo y político como el suizo Jean Ziegler (2006; 2010), quien abogaba por la reconquista de una mentalidad simbólica, por citar tres nombres de procedencia geográfica y cultural diversa, son tan solo una muestra de la diversidad de la «indignación» contemporánea.

Ahora bien, en Europa —específicamente en España— pueden identificarse como ideólogos del movimiento indignado al francés Stéphane Hessel, al español José Luis Sampedro y al uruguayo Eduardo Galeano. Sus

textos y declaraciones, junto a los manifiestos de las agrupaciones ciudadanas, constituyen la base ideológica y argumental —y, por supuesto, textual— del discurso indignado (Cabal 2011). A ellos pueden sumárseles otros autores y proyectos editoriales, incluidos los protagonistas de las acciones sociales, que no han dudado en narrar sus experiencias. A pesar de la diversidad del origen, la formación e incluso los intereses de estos agentes ideológicos, es posible identificar en sus textos y discursos algunos elementos comunes que intentaré destacar de la manera más clara posible.

El movimiento indignado en España, según narran sus promotores, surgió a modo de eco de las revoluciones de los países árabes, sobre todo de la llamada la «revolución del Jazmín» de Túnez, en 2010. Fue la forma de sus convocatorias, y su organización a través de la red, lo que supuso un más claro estímulo para la puesta en marcha de iniciativas paralelas en España, que no tenían mucho que ver, en cuanto a las condiciones del entorno sociopolítico, con lo que estaba ocurriendo en el norte de África y Asia occidental, pero que se identificaban con la idea de buscar «un país mejor para un mundo mejor», como expuso Pablo Gallego en el volumen colectivo *Nosotros, los indignados* (Álvarez *et al.* 2011).

### 2.1. Stéphane Hessel

Y fue al recibirse el eco arabófono cuando en España surgió la voz de Stéphane Hessel, expresada en una obra de 32 páginas publicada en Francia en octubre de 2010 y que en marzo de 2011 ya había sido traducida a 22 lenguas y tenía vendidos 1,5 millones de ejemplares: *Indignez-vous!* La edición en español apareció en febrero de 2011, por lo que su influencia en la gestación misma de la primera iniciativa de la indignación con toda seguridad pudo existir, si bien no podría afirmarse que fuera un elemento detonante o definitivo para la gran marcha convocada para el 15 de mayo de 2011 en Madrid, en Barcelona y en decenas de ciudades españolas. Parece claro, no obstante, que a este opúsculo se le debe la etiqueta que, junto a la de «movimiento 15M», más éxito ha tenido en España para denominar a este fenómeno social: los indignados.

Curiosamente, el original francés *Indignez-vous!* no surgió por iniciativa de su celebrado autor, sino como proyecto concebido y armado por los editores Sylvie Crossman y Jean-Pierre Barou. Este último lo explicaba, con tanta claridad como poca modestia, al periódico español *Público.es* (27-06-2011):



La iniciativa fue nuestra [*Indigène éditions*], lo cual no siempre nos ha sido reconocido. Sin nosotros, el libro no existiría. Creo que Hessel, por su propia voluntad, nunca habría escrito este libro. Fuimos nosotros quienes fuimos a llamar a su puerta. Y fuimos nosotros quienes redactamos el texto a partir de tres entrevistas con Hessel, que después él corrigió y validó antes de su publicación. No me malinterprete: Hessel es la pieza maestra y no quiero minimizar su papel. Pero él mismo reconoce que, sin nuestra intervención, nada habría sido posible.

Cuentan los editores que la idea se les ocurrió al escuchar una conferencia de Hessel en la que se dirigía principalmente a los jóvenes para que no se perdieran las conquistas sociales alcanzadas por la sociedad occidental durante las últimas décadas, exhortación que no era ajena al clima de protestas sociales, algunas de ellas muy violentas, vivido en Francia durante los últimos años. Uno de los primeros reseñantes en español del libro de Hessel, Miguel Izu, concluía su reseña del 11 de enero de 2011 (*Rebellion.org*) preguntándose: «¿Tendrá el mismo éxito Stéphane Hessel en España cuando se edite su libro en castellano?» El tiempo le ha dado la respuesta que ya intuía.

En efecto, Hessel es el autor firmante del gran éxito editorial —y social— titulado, como es sabido, *Indignez-vous!*, que en España la editorial Destino tradujo como *iIndignaos!*, si bien las primeras propuestas de traducción hechas desde España, normalmente en reseñas de prensa y naturalmente con respaldo hispanoamericano, eran *iIndígnese!* o *iIndígnense!*. El original de Hessel, como es habitual en la lengua francesa, se formula con un pronombre de cortesía (*vous*) y con un imperativo que puede dirigirse tanto a un interlocutor/lector singular como a uno plural. En el caso de España, se optó por la traducción con un referente plural y con un tratamiento de cercanía o familiaridad (*vosotros/os*), buscando tal vez un mayor impacto sobre el lector y recurriendo a un tuteo que goza en España de gran generalización y silenciosa aquiescencia. En América, sin embargo, donde la cortesía más formal aún predomina las relaciones sociales, se ha optado por otra traducción, *iIndígnese!*, con referente singular, al que expresamente se le da un tratamiento de *usted*, ya que el título *iIndígnate!* habría sido perfectamente posible, con la variante *iIndígnate!*, de acentuación llana, para las tierras rioplatenses.

El libro de Hessel es bien conocido y no merece ni resúmenes ni reseñas complementarios. Sí nos interesa, en cambio, tratar su contenido desde la

perspectiva del análisis del discurso, con la idea de entender mejor cómo se organiza la comunicación de su mensaje dentro del movimiento indignado. El libro consta de una introducción en la que Hessel apela a su bagaje biográfico e histórico, referido a la Francia de De Gaulle y al primer Consejo Nacional de la Resistencia. Tras ello, vienen sucesivos apartados, titulados «El motivo de la resistencia es la indignación», «Dos visiones de la historia», «La indiferencia: la peor de las actitudes», «Mi indignación a propósito de Palestina», «La no-violencia, el camino que debemos aprender a seguir» y, finalmente, «Por una insurrección pacífica». La obra termina con un ya famoso retruécano, a manera de moraleja positiva: «Crear es resistir, resistir es crear».

El texto entrelaza referencias históricas de Europa, especialmente de la postguerra mundial, con alusiones a diferentes espacios de conflicto (Palestina, Argelia, Estados Unidos, Sudáfrica) y con hechos de la actualidad más cercana, como las reformas educativas de 2008 en Francia, objeto de conflictos con un profesorado que vio reducido su salario como penalización por la desobediencia. Todo ello salpicado de ilustraciones literarias y filosóficas, de Guillaume Apollinaire, Jean Paul Sartre, Marcel Proust, Georg Hegel o Walter Benjamin.

La obra de Hessel permitió acuñar algunos sintagmas que arraigaron en el entramado comunicativo de los indignados: *insurrección pacífica* o *esperanza de la no-violencia*. Ahora bien, en cuanto a su forma discursiva, el texto está construido en torno a dos esferas de percepción claramente identificables: la de un *yo* hablante y la de un *tú* interpelado; en el caso de la versión española, un plural: *vosotros*. El *yo* se manifiesta en las estructuras gramaticales creadas en primera persona del singular o con referencia a ella, comenzando por la declaración de la edad del propio autor: “93 años”. Unos ejemplos:

*Os deseo a todos ...*

*Cuando algo nos indigna, como a mí me indignó el nazismo, ...*

*Cuando intento comprender ...*

*Pero mi optimismo natural [...] me encaminaba más bien a Hegel.*

*No me agunto las ganas de citar el artículo 15 de la Declaración Universal de los Derechos del Hombre.*

*Constato con alegría ...*

*A los jóvenes les digo ...*

*Hoy, mi principal indignación ...*

*Lo sé ...*

*Estoy convencido ...  
He constatado, y no soy el único, ...  
Mi gran pesar ...*

Como es de esperar, en esta esfera del «yo» también se incluye un *yo ampliado o inclusivo*, en forma de primera persona del plural (*nosotros*) que se intercala con los usos en singular. Unas muestras:

*A Jean Moulin le debemos ...  
Estos principios y valores los necesitamos hoy más que nunca.  
Volvemos a encontrarnos con esta corriente de la historia.  
Les decimos con afeción ...*

La mayor parte de estas manifestaciones del *yo* del autor se orientan en tres direcciones discursivas: la declaración, la descripción y la narración, pero no son menos definitorias las que suponen una exhortación: *hagamos un llamamiento; no debemos consentir*. Con todo ello se construye un contexto, un entorno socio-situacional, que fundamenta y justifica la actitud del autor a propósito de los que se interpretan como más que sobrados motivos de indignación. Pero esta esfera del *yo* no tendría pleno sentido si no se construyera en proyección hacia la esfera de un *tú*; o de un *tú* ampliado en *vosotros*. Recordemos que hablamos de una obra editada en España y aparentemente preparada para dirigirse exclusivamente a españoles usuarios de un castellano conservador, puesto que los canarios y muchos andaluces serían más sensibles a un vocativo *ustedes*, no por marcar distancias de cortesía, sino sencillamente porque ese es su pronombre vocativo para la segunda persona del plural, también para la familiaridad, la cercanía o la confianza. El uso del *tú*, ampliado o no, se orienta también en varias direcciones discursivas: la apelación, la exhortación y el imperativo. Las apelaciones y exhortaciones al lector son frecuentes en el texto: *os deseo que ...; Si conocéis a alguien, ...; Al comportaros así, perdéis ...; A los jóvenes les digo: ...* Por otro lado, valgan unos pocos ejemplos de imperativos: *compadecedlo, ayudadle, mirad, ¡Buscad y encontraréis! Y, por supuesto, ¡Indignaos!* El mandato, más moral que coactivo, alterna la modalidad imperativa directa con la impersonalidad:

*No se debe permitir ...  
Debe provocar nuestra indignación ...  
Hay que estar molesto.  
Hay que preferir ...*

*Hay que intentar ...*

*Hay que entender ...*

*No se debe permitir que se acumule mucho odio.*

*Este hecho debería suscitar un compromiso.*

El juego entre el *yo* y el *tú* estructura, pues, el texto de Hessel buscando un efecto de apelación y exhortación a partir de unos hechos descritos o narrados. Sin embargo, esta estructura de esferas discursivas y la contextualización que de ella se deriva no son suficientes para explicar la eficacia alcanzada por el discurso del veterano periodista francés. Aún es necesario descubrir cuáles son las claves que configuran el contenido ideológico de su mensaje y que permiten entender el entorno concebido por el autor, en el que nos quiere situar, para hacer coherente tal contenido con la actitud demandada a los lectores. Estas claves se disponen en dos hilos temáticos (*paradigmas sustantivos* los llamo) que se exigen mutuamente y que constituyen la auténtica columna vertebral de *Indignez-vous!* Uno de esos hilos es el que permite ver los fundamentos de la *indignación*, los elementos que la provocan y que conforman un panorama de negatividad y pesimismo; el segundo hilo dibuja un panorama de *esperanza*, enlazando elementos que la evocan y que la hacen necesaria.

<b>Paradigma sustantivo de la indignación</b>	<b>Paradigma sustantivo de la esperanza</b>
Nazismo	Resistencia
Poder	Responsabilidad
Dividendos	Interés general
Fascismo	Naciones Unidas
Miedo	Libertad
Ricos egoístas	Minoría activa
Ocupación	Descolonización
Totalitarismo	Libertad
Purgas estalinistas	Voluntad de compromiso
Catástrofe	Optimismo
Pequeña élite	Vasto mundo
Indiferencia	Esfuerzo
Hipocresía	Compromiso

Amenazas	Emancipación
Crimen	Insurrección
Totalitarismo	Revulsión
Pueblos colonizados	Derechos
Casas destruidas	Bienestar
Conflicto	Cooperación
Destrucción	Preservación
Prisión	Dignidad
Exasperación	Alegría
Bloqueo	Patriotismo
Odio	Amor
Penurias	Desafíos
Indignación	Esperanza
Violencia	No-violencia
Injusticia	Justicia
Terrorismo	Revoluciones
Guerra	Paz
Opresión	Comprensión
Crisis	Estabilidad
Barbarie	Ética
Retroceso	Organización
Intervención militar	Liberación
Riqueza	Pobreza
Dictadura	Democracia
Fracaso	Relevo
Cólera	Adhesión
Opresores	Oprimidos
Aislamiento	Cultura
Indignación	Esperanza

Estos dos paradigmas, el de la indignación y el de la esperanza, se exigen de modo recíproco y justifican su existencia en la del otro. No tendría razón de ser la exposición de un panorama de desastres si, más allá de las reacciones de indignación, no existiera la esperanza de una regeneración; de igual forma que carecería de justificación un discurso de esperanza sin conocer los elementos que la hacen necesaria. La trabazón de ambas series es tan estrecha que en varios casos se resuelve mediante pares antonímicos plenos, como en *odio / amor, injusticia / justicia, riqueza / pobreza o violento / no-violento*. Por otra parte, estas series léxicas están construidas a base de sustantivos —mayoritarios y decisivos en un análisis de contenido ideológico—, pero debe saberse que también es posible elaborar una relación de adjetivos alineados, muchos de ellos de la misma familia que los sustantivos correspondientes (*indignado, pesimista, fascista, democrático, pacífico, esperanzado, ...*); otros incluso formando pares de antónimos (*violento / no violento, rico / pobre, opresor / oprimido*), junto a algunos muy ilustrativos del «negro» panorama que conduce a la indignación: *destructor, insoportable, horrible, inhabitable, desastroso, ...* Y aún hay que hablar de los verbos, aunque aquí sean menos reveladores y útiles para la caracterización ideológica: en el paradigma de la indignación, *exasperar, desaparecer* y ... , sobre todo, *indignarse* (aparece 4 veces, aparte del título); en el paradigma de la esperanza, *compadecer, ayudar, buscar, conseguir, rechazar, esperar, levantarse* y ..., sobre todo, *comprometerse* (aparece 4 veces). Además, resulta curioso que aparezcan más adjetivos ubicables en el paradigma de la indignación que en el de la esperanza, tal vez por el deseo de calificar adecuadamente la situación de partida; mientras, se encuentran más verbos clasificables en el de la esperanza que en el de la indignación, tal vez por el interés en promover un respuesta activa. En fin, tal dicotomía de paradigmas solidarios, de categorías que se requieren, ha sido vista con toda claridad por Manuel Castells, quien dio a su obra dedicada a esta materia el significativo nombre de *Redes de indignación y de esperanza*. El mensaje que se transmite con todo ello es que, sin indignación, la esperanza es banal y, sin esperanza, la indignación es estéril.

## 2.2. José Luis Sampedro y Eduardo Galeano

La voz española que con más nitidez se ha identificado entre los defensores del movimiento indignado ha sido, sin duda, la del economista José Luis Sampedro Sáez, nonagenario como Hessel. Los discursos más

conocidos de Sampedro a este respecto son el prólogo del libro *¡Indignaos!*, la entrevista divulgada por *Movimiento visual* el 18 de mayo de 2011, aún durante las acampadas, así como su contribución al volumen colectivo *Reacciona* y su participación en otras entrevistas emitidas por la radio y la televisión españolas. Sampedro se presentaba como una voz independiente, autónoma, liberada de compromisos laborales y sociales, concedora del entramado económico desde el que se ha generado la crítica situación internacional.

El discurso «indignado» de Sampedro, tal vez por haberse iniciado con el prólogo al libro de Hessel, refleja algunas de las características formales e ideológicas observadas en el texto del francés, contribuyendo así a la construcción sociocognitiva de un contexto favorable al movimiento indignado. En el prólogo de la obra de Hessel, también es posible construir una paradigma de la indignación (*culpables de la crisis, paraísos fiscales, poder, dividendos, carrera por el dinero, culpabilidad, totalitarismo, huracán destructor, engaño propagandístico, consumismo voraz, distracción mediática, atropello, indignación*) opuesto a un paradigma de la esperanza (*voluntad de compromiso, avances, estado del bienestar, logros democráticos, valores éticos* y, por supuesto, *insurrección pacífica, sin violencia*). Los paradigmas se entrecruzan y se requieren, pero, de nuevo, se proyecta la experiencia de un *yo* (ampliado o no) sobre la llamada a un *tú* (ampliado o no) al que se apela y se exhorta. Por eso la forma lingüística también incorpora como rasgos definitorios la pregunta (retórica o no), el apelativo y el imperativo: *luchad; negaos; indignaos*.

No por casualidad, el doble eje que configuran los textos fundacionales del movimiento #15M (de la indignación a la esperanza / de la descripción a la exhortación) se reproduce en los discursos orales de Sampedro, como en la entrevista aparecida en *Movimiento visual*. Allí habla de la existencia de un sistema disgregado y corrompido, de la manipulación de la gente, del dominio del poder sobre los medios de información, del hundimiento del sistema financiero y capitalista y de la sustitución, en Occidente, de los «valores» por los «intereses». Pero, casi en cada afirmación, el contrapunto emerge con claridad y se habla de «hacerse más persona», de ayudar a cambiar las cosas, de tomar conciencia de la realidad, de seguir adelante. —«Se ha degradado el sistema» —dice Sampedro; para añadir acto seguido: —«Hay alternativa. Seguro». Ese mismo juego dialéctico de tesis-antítesis lo prolongó Sampedro en una entrevista que Iñaki Gabilondo le hizo para «Canal +. España», titulada «Iñaki & José Luis Sampedro» (12/08/2012),

donde nuestro autor opone «productividad – competitividad – innovación» a la tríada renovadora «vitalidad – cooperación – recreación», aun cuando el contenido del programa iba más dirigido hacia su persona que hacia una valoración del movimiento indignado. Sampedro estableció allí un claro paralelismo entre el movimiento del mayo del 68, con su más conocido eslogan («que paren el mundo, que me apeo»), y los movimientos indignados.

Igualmente interesante, aunque no muy distinto en cuanto a su fuerza perlocutiva o efecto buscado, es el texto que, con el título «Debajo de la alfombra», Sampedro aportó al volumen colectivo *Reacciona* (Artal 2011). A la hora de calificar el pasado reciente y el devenir contemporáneo de la humanidad, el autor lo hace de una forma muy expresiva: «Diez en técnica y cero en humanismo». He aquí otro descubrimiento «debajo de la alfombra» que revela una maraña de injusticias, irracionalidades y hasta delitos, impunes o encubiertos, dejando la impresión desoladora de la brillantez de la técnica y la ciencia se debe a un organismo social vivo, cuyas contradicciones internas lo están descomponiendo. De ahí la sensación de ocaso y de barbarie, entre cuyos componentes está la crisis. Ahora bien, el ocaso no es el fin de la historia, sino del sistema (Artal 2011: 19). La palabra *ocaso* es clave en la interpretación de José Luis Sampedro: por un lado como consecuencia de los aspectos negativos escondidos debajo de la alfombra de las sociedades occidentales; por otro lado, como punto de inflexión para la acción, en busca de otro mundo que no solo es posible, sino seguro.

Eduardo Galeano, autor uruguayo representante destacado del pensamiento crítico iberoamericano, tuvo una presencia casi circunstancial durante los acontecimientos sociales del mes de mayo de 2011. En ruta hacia el cumplimiento de otras obligaciones, Galeano fue entrevistado de modo informal durante la primera visita que realizó a la acampada de la ciudad de Barcelona, lo que dio lugar a otras entrevistas y declaraciones, incluida la realizada en los estudios de la emisora catalana TV3, en junio de 2011. El discurso de Galeano no responde a una estructura tan nítida como la que se ha descrito para Hessel y Sampedro, tal vez porque, en relación con el movimiento indignado, está construido sobre la oralidad y a partir de una estructura de entrevistas.

Eduardo Galeano es un pensador que lleva tiempo prodigándose en sus críticas a la peor cara de las sociedades del primer mundo, por lo que sus argumentos más y mejor probados han venido a revitalizarse en el contexto de la indignación, hasta casi adquirir la naturaleza de advertencia desatendida. Aun así, la estrategia discursiva consistente en una descripción



de calamidades, generadora de una indignación que ha de mitigarse con en el entusiasmo de la esperanza, vuelve a aparecer con nitidez en Galeano y a evidenciar la solidez de un contexto que ayuda a interpretar y justificar la esencia del movimiento indignado. De esta manera, el «mundo infame» de las agencias de calificación —cita expresamente, por lo paradójico del nombre, *Standard & Poor's*—, los «tecnócratas de cuarta», los *ignorantes*, el «mundo que recompensa a sus arruinadores», el «mundo de mierda», el «mundo al revés» o los «políticos mentirosos» ceden su crédito a los jóvenes, a la democracia —«secuestrada por los banqueros»—, a la *energía de dignidad*, al *entusiasmo* —etimológicamente ‘posesión divina’— a la vida que «vale la pena» y a la «vitamina E, de entusiasmo», que es la que se ha generado, según Galeano, en los movimientos sociales de mayo de 2011.

Durante la entrevista que el programa «Singulars» le hizo a Galeano en TV3 (2011), volvió a referirse, casi literalmente, al «entusiasmo» y afloraron de nuevo unidades sustantivas del paradigma de la esperanza: *protesta pacífica, solidariamente, convicción, voces capaces de decir «basta», lecciones de dignidad, movimiento nuevo (lindo), derecho a soñar, derecho al delirio, utopía*; unidades que se entrecruzan con las del paradigma que dibuja una realidad desoladora: *tecnocracia, enfermos de «exitismo», crímenes de lesa humanidad, dictadura financiera*, «el mercado tiene mal humor». Y la dicotomía se explicita en esta afirmación literal: «el mundo está dividido entre los indignos y los indignados». La diferencia de este discurso respecto de los de Hessel y Sampedro es la menor fuerza de su llamada a la acción.

### 3. LOS MANIFIESTOS

La palabra *manifiesto*, como sustantivo, tiene una historia al menos tricentenaria en las lenguas de origen europeo. Las referencias más antiguas dan a este sustantivo el significado genérico de 'aviso, anuncio, declaración'. En italiano se registra desde el siglo xvii con el significado de 'papel impreso expuesto al público para dar hacer avisos o dar publicidad a un asunto'. En francés hay referencias de *manifeste* como 'escrito público destinado a hacer conocer ideas nuevas sobre un dominio determinado' desde 1623 (CNRTL, s.v.), si bien desde el siglo xvi designaba al documento que redacta un responsable político para explicar su conducta. En inglés se documenta el

italianismo *manifesto* desde 1620, con el significado de 'declaración pública donde se explican acciones pasadas y se anuncian los motivos de acciones futuras' (*Online etymology Dictionary*, s.v. *manifest*).

La lengua española, como refleja el *Tesoro de la lenguas castellana o española* de Sebastián de Covarrubias, cuenta con el sustantivo *manifesto* como 'manifestación, declaración' desde antes de 1611. Asimismo desde la primera edición del diccionario académico, el llamado *Diccionario de Autoridades* (1734), se recoge el significado de 'escrito con el que se justifica o aclara alguna cosa al público', ilustrándose con un texto de la *Crónica de la orden de San Francisco*, de fray Damián Cornejo, publicada entre 1682 y 1698: «se imprimió en un manifiesto, hecho en cabeza de toda la Religión, y se divulgó por toda Europa». En la edición de 1803, la Academia retocó su definición —«Escrito con que se justifica o manifiesta alguna cosa»—, que permaneció inalterada hasta la edición de 1925, en la que se incluía la acepción que ha conservado vigencia hasta el siglo xxi: «Escrito en que se hace pública declaración de doctrinas o propósitos de interés general».

¿Qué es lo que propició la modificación de la definición académica, llevándola al terreno de la ideología (declaración de doctrinas) y de la política (propósito de interés general)? Probablemente la resonancia alcanzada por un manifiesto de 1848, quizá el más famoso «manifiesto» de la historia de Occidente, tras las 95 tesis de Lutero: el *Manifiesto del partido comunista*, de Carlos Marx y Federico Engels, aquel que comenzaba diciendo «Un fantasma recorre Europa: el fantasma del comunismo». La amplia e intensa repercusión del manifiesto comunista propició que las organizaciones de izquierdas comenzaran a hacer uso frecuente de este tipo de declaración, a la vez que se convertía en vehículo de expresión de movimientos heterodoxos o innovadores de todo tipo: artísticos, políticos, ideológicos, sociales, ... En la difusión internacional de este medio de expresión se ha ido produciendo una progresiva pérdida del deslinde entre los «manifiestos» propiamente dichos, los «llamamientos» y las «declaraciones», de manera que muchos documentos de cualquiera de estos tipos bien podrían intercambiar con otros su etiqueta clasificadora. Actualmente, manteniéndose la amplitud semántica del término, parecen reservarse al «manifiesto» las características de declarar doctrinas y propósitos de interés general, generalmente de tipo sociopolítico, la de describir acciones pasadas para anunciar otras futuras y la de hacerse público a través de los medios de comunicación social.

Los manifiestos son, sin duda, textos clave en los procesos comunicativos del movimiento indignado, entre otras razones porque son capaces de cumplir varias funciones: describir unos hechos considerados como inaceptables; declarar un posicionamiento ante una realidad; establecer una conveniencia; acreditar a sus firmantes; exhortar a la acción. Los manifiestos aspiran a reflejar unas pautas esenciales, entre las que se halla la diversidad. Esta se evidencia, desde el punto de vista idiomático, en la publicación de textos en diversas lenguas, con actitud ecuménica. Por eso los manifiestos indignados de España se dan a conocer en español, gallego, catalán, euskera o inglés; y también en un asturiano diferenciado de un *llionés*. En cuanto a su contenido, los manifiestos indignados pueden interpretarse de una doble forma: como vehículos justificativos de unas acciones sociales de gran repercusión mediática y como instrumentos argumentativos ante un *statu quo* político-económico indeseable y contra los movimientos que han intentado minimizar la incidencia social de la indignación. Esta interpretación convierte a los manifiestos en elementos clave en la gestión de las acreditaciones ofrecidas y de las conveniencias denunciadas por sus autores (Potter 1998), tanto para la presentación pública y el desarrollo interno del propio movimiento indignado, como para su defensa ante los discursos contrarios. Además, los argumentos exhibidos vienen acompañados de una exhortación a la acción «indignada» o a otro tipo de movilizaciones más concretas.

Para conocer más de cerca cómo están contruidos los manifiestos del movimiento indignado, utilizaremos dos ejemplos: el manifiesto titulado «Yo #soy15m» y el manifiesto «Democracia real ya» promovido por la plataforma del mismo nombre. Hemos elegido el primero por tener una amplísima difusión en la red —el buscador «Google»<sup>®</sup> ofrecía más de 15000 resultados bajo ese enunciado en enero de 2013—; y el segundo, por proceder de la agrupación convocante de las manifestaciones de mayo de 2011 en España. Uno y otro, al mismo tiempo, ofrecen ciertos rasgos diferenciadores, que enriquecen el análisis.

#### MANIFIESTO “YO #SOY15M”

Como parte del #15m me declaro una persona pacífica y condeno radicalmente todo tipo de violencia: la de los violentos infiltrados en nuestras manifestaciones, y la del Estado, que ha causado más dolor y heridos. Además, condeno la manipulación mediática que enfatiza la información sesgada, parcial o errónea con el propósito de demonizar a los ciudadanos.

Si me manifiesto en la calle es porque:

Mi participación como ciudadano se ha reducido a votar a listas cerradas cada cuatro años para ver cómo los representantes de los ciudadanos no respetan lo prometido en su programa.

Se hacen leyes a favor de grupos de interés en vez de hacerlas a favor del conjunto de la sociedad.

Se invierten recursos públicos para ayudar a minorías poderosas, y no a quienes están pasando situaciones desesperadas ocasionadas por la especulación financiera.

Los grandes partidos están más preocupados por mantener su poder que por ofrecer soluciones para superar esta crisis histórica.

Está a punto de firmarse un «Pacto del Euro» que consiste fundamentalmente en medidas para reducir la inversión pública en servicios esenciales.

Desde diferentes órganos del estado se ha insultado a los ciudadanos, e incluso se ha justificado el recurso a la violencia contra manifestantes pacíficos.

Como parte del #15m, acepto y respeto la diversidad ideológica del movimiento. Cuando participo en una manifestación no reclamo un régimen o una ideología en concreto, ni un modelo social no democrático, ni la eliminación de los partidos o los parlamentos. Lo que reclamo es una democracia mejor y más humana que, entre otras medidas, necesita urgentemente:

Cambios en la Ley Electoral para permitir una mejor y más directa representación de los ciudadanos en los parlamentos y una mayor participación ciudadana en las decisiones importantes.

Aprobación de una Ley de Transparencia y Acceso a la Información Pública para obligar a la publicación en formatos adecuados y reutilizables de todos los gastos, decisiones y reuniones con grupos de presión por parte de funcionarios y cargos públicos.

Tolerancia cero a la corrupción de candidatos y cargos públicos, y controles ciudadanos para la exigencia de responsabilidad política.

Separación clara, real y efectiva de los poderes del estado.

Control fiscal efectivo de grandes fortunas y operaciones financieras; eliminación de privilegios fiscales a cargos electos.

Políticas encaminadas a solucionar de forma efectiva los problemas hipotecarios y de vivienda.

Servicios públicos de calidad, fundamentalmente salud, justicia y educación.

Eliminación de las leyes que permiten el control administrativo de Internet. La red ha demostrado ser esencial para la libertad de expresión y para responder al peligro de manipulación mediática.

Por todas estas razones volveré a salir pacíficamente a la calle el 19 de junio, #19j.

Si estás de acuerdo, aprópiate del texto y divúlgalo.

El texto del primer manifiesto alude en su título al 15 de mayo de 2011 y está redactado en primera persona del singular. El texto busca la firma personal o colectiva de todos aquellos que deseen adherirse a su contenido, trasladando la convicción de los redactores a los posibles suscriptores, sin posibilidad de que pueda incluirse modificación ni matización alguna. La perspectiva desde la que está redactado es, por lo tanto, la de un *yo* que solo se rompe en la última oración, cuando se incluye una llamada metatextual para que el lector contribuya a la divulgación del documento.

Es significativo que el primer verbo del texto sea *declaro*, hecho que establece la finalidad principal del mensaje, a la vez que anuncia la primacía discursiva de las formas verbales de tipo declarativo (*acepto, condeno, manifiesto, reclamo, ...*). El manifiesto está organizado en una primera parte declarativa, seguida de dos partes de estructura enumerativa –una para explicar las razones de la declaración y una segunda para exponer reclamaciones o reivindicaciones– y finaliza en el anuncio de una acción futura. Como se observa, esta configuración es claramente la que se daba en los *manifestos* de la Inglaterra del siglo xvii: explicar hechos pasados y anunciar acciones futuras.

Desde el punto de vista del léxico, como reflejo del ideario transmitido por los autores fundacionales y como aportación a la construcción de un contexto acorde con el posicionamiento indignado, es interesante identificar, como en los textos de Hessel o Sampedro, dos paradigmas perfectamente diferenciados y enfrentados: el de la indignación y el de la esperanza. El primero vertebrado por la relación de causas que llevan al manifiesto de indignación y está constituido por expresiones sustantivas como *listas cerradas, minorías poderosas, situaciones desesperadas, especulación financiera, crisis histórica, violencia*. El segundo paradigma permite elaborar una enumeración de reclamaciones y fijar unos objetivos cuyo cumplimiento se espera: *diversidad ideológica, democracia mejor, participación ciudadana,*

*responsabilidad política, libertad de expresión.* La carga positiva de estas expresiones se refuerza cuando, en algunos casos, se oponen en las mismas frases a sus antónimas: *corrupción de candidatos / controles ciudadanos; privilegios fiscales / control fiscal; manipulación mediática / libertad de expresión.* Por último, en el manifiesto «Yo #soy 15M» aparece tres veces una palabra esencial en la acción de los indignados, según sus propios textos: el adjetivo «pacífico» y su derivado adverbial «pacíficamente»: por eso se encuentra en la declaración inicial y en el anuncio que cierra el manifiesto: «me declaro una persona pacífica»; «volveré a salir pacíficamente.» Este elemento parentético, que abre y cierra el texto, es una prueba más de la cohesión formal del manifiesto y de su coherencia discursiva.

El segundo de los manifiestos muestra algunos rasgos comunes con el recién analizado, pero también algunas diferencias importantes, especialmente en lo que a su perspectiva y a su estructura se refiere.

#### **MANIFIESTO “DEMOCRACIA REAL YA”**

Somos personas normales y corrientes. Somos como tú: gente que se levanta por las mañanas para estudiar, para trabajar o para buscar trabajo, gente que tiene familia y amigos. Gente que trabaja duro todos los días para vivir y dar un futuro mejor a los que nos rodean.

Unos nos consideramos más progresistas, otros más conservadores. Unos creyentes, otros no. Unos tenemos ideologías bien definidas, otros nos consideramos apolíticos... Pero todos estamos preocupados e indignados por el panorama político, económico y social que vemos a nuestro alrededor. Por la corrupción de los políticos, empresarios, banqueros... Por la indefensión del ciudadano de a pie.

Esta situación nos hace daño a todos diariamente. Pero si todos nos unimos, podemos cambiarla. Es hora de ponerse en movimiento, hora de construir entre todos una sociedad mejor. Por ello sostenemos firmemente lo siguiente:

Las prioridades de toda sociedad avanzada han de ser la igualdad, el progreso, la solidaridad, el libre acceso a la cultura, la sostenibilidad ecológica y el desarrollo, el bienestar y la felicidad de las personas.

Existen unos derechos básicos que deberían estar cubiertos en estas sociedades: derecho a la vivienda, al trabajo, a la cultura, a la salud, a la educación, a la participación política, al libre desarrollo personal, y derecho al consumo de los bienes necesarios para una vida sana y feliz.

El actual funcionamiento de nuestro sistema económico y gubernamental no atiende a estas prioridades y es un obstáculo para el progreso de la humanidad.

La democracia parte del pueblo (demos=pueblo; cracia=gobierno) así que el gobierno debe ser del pueblo. Sin embargo, en este país la mayor parte de la clase política ni siquiera nos escucha. Sus funciones deberían ser la de llevar nuestra voz a las instituciones, facilitando la participación política ciudadana mediante cauces directos y procurando el mayor beneficio para el grueso de la sociedad, no la de enriquecerse y medrar a nuestra costa, atendiendo tan sólo a los dictados de los grandes poderes económicos y aferrándose al poder a través de una dictadura partidocrática encabezada por las inamovibles siglas del PPSOE.

El ansia y acumulación de poder en unos pocos genera desigualdad, crispación e injusticia, lo cual conduce a la violencia, que rechazamos. El obsoleto y antinatural modelo económico vigente bloquea la maquinaria social en una espiral que se consume a sí misma enriqueciendo a unos pocos y sumiendo en la pobreza y la escasez al resto. Hasta el colapso. La voluntad y fin del sistema es la acumulación de dinero, primándola por encima de la eficacia y el bienestar de la sociedad. Despilfarrando recursos, destruyendo el planeta, generando desempleo y consumidores infelices.

Los ciudadanos formamos parte del engranaje de una máquina destinada a enriquecer a una minoría que no sabe ni de nuestras necesidades. Somos anónimos, pero sin nosotros nada de esto existiría, pues nosotros movemos el mundo.

Si como sociedad aprendemos a no fiar nuestro futuro a una abstracta rentabilidad económica que nunca redunde en beneficio de la mayoría, podremos eliminar los abusos y carencias que todos sufrimos.

Es necesaria una Revolución Ética. Hemos puesto el dinero por encima del Ser Humano y tenemos que ponerlo a nuestro servicio. Somos personas, no productos del mercado. No soy sólo lo que compro, por qué lo compro y a quién se lo compro.

Por todo lo anterior, estoy indignado.

Creo que puedo cambiarlo.

Creo que puedo ayudar.

Sé que unidos podemos.

Sal con nosotros. Es tu derecho.



Entre lo común, resulta inevitable redescubrir el doble paradigma que explica el tránsito de la indignación a la esperanza, por medio de una acción reclamada. En este caso, sin embargo, los sustantivos —o frases nominales— aparecen en su mayor parte seriados dentro de un mismo párrafo o incluso de una misma frase. Comienzan las series fijando un horizonte de esperanza constituido por sustantivos, casi todos abstractos, como *igualdad, progreso, solidaridad, sostenibilidad, desarrollo, bienestar, felicidad*. El contrapunto llega al enumerar las consecuencias del mal uso del poder: *desigualdad, crispación, injusticia, violencia, pobreza, escasez*, hasta un rotundo ... *colapso*. Esta retahíla de sustantivos negativos también aparece en los textos ensayísticos de los indignados, como en este fragmento de Pablo Gallego (2011): «Las seis causas de la crisis en España han sido, básicamente, la corrupción, el gasto público, la especulación inmobiliaria, la privatización de las cajas, la inacción de la ciudadanía y el abismo entre los políticos y los ciudadanos» (*Nosotros, los indignados*, p. 28). El entorno situacional queda, pues, deslindado de una forma clara y coincidente con los textos fundacionales y con el manifiesto «Yo #soy15M».

Ahora bien, si el primer manifiesto ofrecía una estructura equilibrada, desde una perspectiva homogénea, dotada, por tanto, de una cohesión robusta, este segundo manifiesto ofrece una diversidad de perspectivas —una polifonía, podríamos decir— que resta elegancia al discurso —no necesariamente eficacia comunicativa— por querer atender a fines diferentes con una misma declaración. El manifiesto «Democracia real ya», sin coma ante el adverbio, muestra claramente una primera perspectiva del *nosotros*, no siempre intercambiable con la de un *yo* ampliado, como se aprecia en el mismo arranque: «Somos personas normales y corrientes». Existe una segunda perspectiva del *yo*, referida a una primera persona como firmante, utilizada en la parte final del texto, que concluye introduciendo una nueva perspectiva, la de un *tú* exhortado a la acción de *salir*: «Sal con nosotros». Y a ello debe añadirse una cuarta perspectiva externa (*él / ella / ellos*) que describe, con tiempos de presente, una realidad verificable y unos hechos que fundamentan, cuando no justifican, las perspectivas de las primeras y la segunda persona, para la cual se recurre a verbos que denotan existencia (*existir, ser, hacer, generar, ...*). Cuando la perspectiva es la del *yo*, el protagonismo es de los verbos declarativos y de creencia (*considerar, sostener, creer, saber*).



#### 4. CONCLUSIÓN

Los movimientos indignados tienen en los textos fundacionales de sus ideólogos y en sus manifiestos la base discursiva a partir de la cual se ha desarrollado buena parte de las estrategias sociales y comunicativas de sus simpatizantes. En todos ellos pueden observarse elementos comunes, que revelan la contraposición y complementariedad entre unos sentimientos de indignación y unos atisbos de esperanza. El análisis del discurso nos ha permitido descubrir unas estrategias comunicativas desplegadas por medio de unidades y estructuras lingüísticas.

Ahora bien, el estudio de la lengua y la comunicación en el mundo de los indignados no se agota con un somero análisis de sus discursos fundacionales. Se hace necesario contrastar estos discursos con los asociados a la anti-indignación, contraponer estrategias comunicativas, asociarlas a intereses que trascienden el mundo de la expresión lingüística. Y aún es posible ir más allá porque los lenguajes indignados se manifiestan de forma oral y escrita, personal y colectivamente, a través de otros medios y registros. Valgan como simples ejemplos los procesos comunicativos que se despliegan en las asambleas públicas, la exhibición de eslóganes y consignas que acompañan cada acción social o el desarrollo de páginas electrónicas y portales asociados directamente a la indignación. Todo ello deberá ser objeto de estudio en su momento y ser analizado desde un modelo teórico lo suficientemente flexible y amplio como para poder dar cuenta de tan complejas realidades.

## REFERENCES

- Álvarez, Klaudia, Pablo Gallego, Fabio Gándara y Óscar Rivas (2011): *Nosotros, los indignados. Las voces comprometidas del #15-M*. Barcelona: Destino.
- Artal, Rosa M.<sup>a</sup>. (coord.) (2011): *Reacciona*. Madrid: Aguilar.
- Cabal, Fernando (2011): *¡Indignados! 15M*. Ediciones Mandala.
- Castells, Manuel (2009): *Comunicación y poder*. Madrid: Alianza.
- (2012): *Redes de indignación y de esperanza*. Madrid: Alianza.
- CNRTL [Centre National de Ressources Textuelles et Lexicales]. *Trésor de la Langue Française informatisé*.  
<http://www.cnrtl.fr/etymologie/manifeste>.  
[Consultado: 29/10/2013]
- Covarrubias, Sebastián de (1611): *Tesoro de la lengua castellana o española*. Madrid. Ed. de I. Arellano y R. Zafra. Madrid / Frankfurt: Iberoamericana / Vervuert; Universidad de Navarra. 2006
- Galeano, Eduardo (2011): Entrevista. *Singulars*. Programa de TV3. 23 de mayo de 2011.  
<http://www.tv3.cat/videos/3541530>.
- Hessel, Stéphane (2010): *Indignez-vous!*. Paris: Indigène editions. Trad. al esp. de María Belvis Martínez García: *Indignaos*. Barcelona: Destino. 2011. Prólogo de José Luis Sampedro.
- Martín Corvillo, José Manuel (2011). *15M. Anatomía lingüística de la indignación*. Valencia: Universitat de València.
- Martín Rojo, Luisa (2013): “Paisajes lingüísticos de indignación. Prácticas comunicativas para tomar las plazas”. *Anuari del conflicte Social 2012*, pp. 275-302.  
<http://revistes.ub.edu/index.php/ACS/article/download/6276/8020>.  
[Consultado: 29/10/2013].

- Moreno Fernández, Francisco (2012): *Sociolingüística cognitiva. Propositiones, escolios y debates*. Madrid – Frankfurt: Iberoamericana – Vervuert.
- Online Etymology Dictionary.  
<http://www.etymonline.com/>.  
 [Consultado: 29/10/2013].
- Potter, Jonathan (1998): *La representación de la realidad. Discurso, retórica y construcción social*. Barcelona: Paidós.
- Pujante, David y Esperanza Morales López (2013): “Discurso (discurso político), constructivismo y retórica: los eslóganes del 15-M”. *Language, Discourse, & Society* 2/2, pp. 33-59.  
<http://www.language-and-society.org/journal/issues.html>.  
 [Consultado: 29/10/2013].
- Real Academia Española (1726-1739): *Diccionario de la lengua castellana*. Edición facsímil. *Diccionario de autoridades* Madrid: Gredos, 1963.
- Sábato, Ernesto (2000): *La resistencia*. Barcelona: Seix Barral
- Sampedro, José Luis (2011): “Debajo de la alfombra”. En Rosa M. Artal (coord.), *Reacciona*. Madrid: Aguilar, pp. 13-24.
- Schiffrin, Deborah (1994): *Approaches to Discourse*. Oxford: Blackwell.
- Van Dijk, Teun A. (2011): *Sociedad y discurso*. Barcelona: Gedisa.  
 – (comp.) (2008): *El discurso como interacción oral*. Barcelona: Gedisa.
- Ziegler, Jean (2006): *El imperio de la vergüenza*. Madrid: Taurus.  
 – (2010): *El odio a occidente*. Madrid: Península.



## **Discurso e identidad: grupos originarios de México en Carolina del Norte**

Liliana Paredes

*Duke University*

D. Kyle Danielson

*The University of British Columbia*

**Abstract:** The arrival of Mexican immigrants to North Carolina brought a new social and ethnic profile to the city of Durham. However, what has formerly been seen principally as a Spanish-speaking presence now ceases to be a monochromatic product when one notices that a significant number of members of the population represent an ethnic, cultural, and linguistic diversity that is all but invisible to the untrained eye. This work examines the discursive manifestations of immigrants of indigenous origin whose second language is Spanish, and whose interactive sociolinguistic space is reconfigured when they must incorporate other ethnicities and social groups into their new interactive context. This research seeks to respond to the question of how speakers from different social and ethno linguistic groups manifest cultural and ethnic differences, as well as conflicts and social inequities in their discourse within their own group. In this article we posit that through their discourse, speakers from different ethnic origins negotiate their identity and at the same time express their perceptions in terms of relationships of power and dominance. This research is based on recorded conversations with speakers of Chatino, Nahuatl, and Otomí origin.

**Keywords:** Identity, indigenous languages, discourse, immigration.



## 1. INTRODUCCIÓN

La región del Piedmont en Carolina del Norte se ha convertido en una zona de alto crecimiento de comunidades migrantes de México en Estados Unidos. Además, ésta es un área que ha atraído a poblaciones de hablantes de lenguas originarias de México que coexisten al lado de hablantes de español como primera lengua y, por su puesto, al lado de hablantes de inglés<sup>1</sup>. Todo esto hace que el panorama socio y etnolingüístico en Durham sea un espacio de interacción donde la identidad se negocia de acuerdo a expectativas formadas por una historia de relaciones de poder y discriminación. La supuesta homogeneidad del grupo unificadamente llamado latino es desafiada por esta realidad.

En este artículo se presenta y discute esta situación sociolingüística desde una perspectiva que examina el discurso de hablantes que representan dos lados de la nueva situación. Por un lado, observamos lo expresado por los tradicionalmente llamados latinos, hispanohablantes monolingües, es decir, personas que aprendieron español como lengua materna y no de manera bilingüe. Por otro lado, estudiamos el discurso de los hablantes de lenguas originarias bilingües en español cuya presencia en Durham es evidente.

En cuanto a identidad y lengua, el punto de partida que asumimos es que la nueva situación sociolingüística en la que se encuentran los hablantes de lenguas originarias de México les permite replantear y negociar su identidad en oposición a lo latino, y en respuesta a una reproducción de la historia de discriminación y marginación sufrida en México.

Esta negociación y reconstrucción de identidad cobran importancia además, si se entiende que los procesos migratorios incluyen procesos de aculturación en el nuevo ámbito social en el que se insertan. En este trabajo se presupone una definición de aculturación dinámica y fluida (sin puntos finales, pero con estadios variables) (Padilla y Pérez 2003) que permite explicar el proceso que vivían los hablantes de lenguas originarias en Durham en el 2009. En el modelo presentado por Padilla y Pérez la aculturación es un proceso en el que interactúan tanto los inmigrantes como la población que los recibe en el nuevo espacio compartido. Así, según los autores hay cuatro pilares que sostienen los procesos de aculturación: cognición social,

---

<sup>1</sup> Desde el 2009 ha habido un incremento en la presencia de hablantes francófonos y de hablantes de árabe gracias a programas de refugiados. Todas estas lenguas coexisten en el mismo espacio geográfico.

identidad social, conciencia cultural (tanto de la cultura de origen como de la o las culturas con las que establecen relación) y el estigma<sup>2</sup>. El estigma tiene un peso importante en cómo los inmigrantes perciben su situación social.

## 2. EL PANORAMA SOCIOLINGÜÍSTICO DE DURHAM

Aunque el número exacto de hablantes de lenguas originarias en Durham no está definido dado que muchos tienen el estatus de inmigrantes sin autorización de estar en este país, es claro que en esta región hay una presencia significativa de hablantes de Otomí, Náhuatl, Chatino, Mixteco además de lenguas originarias de Centro América. La investigación realizada se llevó a cabo a lo largo de diez meses en los años 2008 y 2009 y consiste en entrevistas sociolingüísticas con hablantes de Náhuatl, Chatino y Otomí, procedentes de los estados mexicanos de Puebla y Oaxaca. Con la excepción de una pareja Nahua (que era la más joven de todos los participantes), los hablantes entrevistados indicaron que su lengua materna era la lengua originaria la cual además usaban como su medio de comunicación primario diariamente. Un dato importante respecto de esta pareja es que eran los más jóvenes de los hablantes con quienes pudimos entablar entrevistas. Además todos los hablantes hablaban español con una variedad de nivel de proficiencia.

### 2.1. Comunidad de habla

Es importante indicar, que nuestro acercamiento a la comunidad de habla de lenguas originarias no fue ni fácil ni inmediato. Particularmente hace 8 ó 9 años, un segmento importante de esta población era poco conocida. Básicamente las comunidades de habla originaria eran poblaciones escondidas por al menos dos motivos. El primero, la falta de documentación para trabajar en el país; el segundo, la falta de integración con la comunidad hispanohablante más grande. Metodológicamente, esto implicó encontrar un número reducido de personas dispuestas a realizar entrevistas orales. Por este motivo, presentamos solamente un panorama de lo encontrado en aquel momento. En este contexto, no es de sorprender encontrar que otro rasgo que caracteriza la recolección de datos es la distancia que toman algunos de los hablantes de los entrevistadores. Esto resulta en entrevistas breves y con material lingüístico limitado.

<sup>2</sup> Nuestra traducción de: 'Social Cognition, Social Identity, Cultural Awareness And Stigma'.

El cuadro 1 presenta la descripción demográfica de los hablantes con los que pudimos realizar alguna entrevista. Cabe anotar que en este grupo, el dominio de español estaba marcado por las variables género y edad. A más joven, mayor dominio de español. Por otro lado, hablantes mayores y hablantes mujeres expresaban menor dominio de español. Es interesante observar también el rango de nivel de educación declarada.

Código de Hablante	Edad	Género	Tiempo en EE.UU.	Lengua originaria	Nivel de educación <sup>3</sup>
M.B	47	M	30	Otomí	primaria
M.E.B.	46	F	8	Otomí	primaria
R.A.	22	M	6	Otomí	primaria
B.H.	33	M	?	Otomí	universidad
G.E.	35	F	13	Otomí	secundaria
F.E.	31	M	13	Náhuatl	ninguna
R.E.	23	F	2	Náhuatl	preparatoria
R.C.	29	F	8	Chatino	primaria

**Cuadro 1.** Datos demográficos de los hablantes.

También se condujeron entrevistas con hablantes cuyo español fue adquirido de manera monolingüe<sup>4</sup>. Este grupo de hablantes incluye personas de México y otros países que tienen diferentes funciones en la comunidad. Algunas de estas personas trabajan en organizaciones cuya misión es atender necesidades de poblaciones vulnerables en la ciudad de Durham. El objetivo de estas entrevistas fue establecer una comparación o contraste en cuanto a actitudes hacia los grupos originarios hablantes de español como segunda lengua. Estas entrevistas se llevaron a cabo en español y en inglés y se condujeron de manera informal en diferentes contextos situacionales.

## 2.2. Migración y Asentamiento

El arribo de las lenguas originarias de los pueblos de México a Durham no es directo. En todos los casos explorados los hablantes de estas lenguas

<sup>3</sup> Los niveles de educación en México son: primaria (6-12), secundaria (12-15) y preparatoria (15-18).

<sup>4</sup> Nos referimos a los hablantes que no tienen en su fondo lingüístico una lengua originaria además del español. Sin embargo éstos tenían un nivel de conocimiento de inglés, lo cual hace que su variedad de español fuera una variedad de contacto.



hacen escalas en diferentes zonas urbanas en México que incluyen el Distrito Federal, y en algunos casos algunas ciudades en los Estados Unidos. Además, la inmigración a regiones como la del Piedmont de Carolina del Norte ocurre frecuentemente por la presencia de contactos sociales que precedían a la inmigración.

Las relaciones laborales que establecen muchos de los hablantes de lenguas originarias en Durham están asociadas con los negocios de restaurantes y la agricultura. Varios de los hablantes que formaron parte de este estudio indicaron además una conexión estrecha con el trabajo en restaurantes chinos (ver Cruz 2006). La distinción laboral también está dada por la diferencia de géneros. Las mujeres se dedican a labores domésticas ya sea en casas particulares, hoteles o empresas. Sin embargo, cabe anotar que algunos de los hombres de origen otomí, particularmente aquellos que tienen un tipo de presencia social en su grupo no se limitan a trabajar para otros. Estos pueden ser trabajadores independientes o incluso pequeños empresarios con actividades económicas que van desde la tienda de abarrotes, el negocio de pintura, a una tienda de medicina natural. Lo más interesante desde el punto de vista socio-económico es que estos negocios contratan mayormente a personal de origen amerindio, de su propia etnia.

En términos de su ubicación geográfica en Durham, las poblaciones latinas en general, y las de origen indígena en particular, no están concentradas en un solo punto de la ciudad. Entonces, a pesar de las redes sociales establecidas por medio de los negocios o del trabajo, el sentido de comunidad está geográficamente descentralizado. En el 2009 cuando se le pregunta al individuo sobre el sentido de comunidad, la respuesta generalizada es la ausencia de un sentimiento comunitario con respecto a grupos de hablantes de lenguas amerindias. Aunque los hablantes entrevistados reconocen la presencia de más hablantes de sus propias lenguas, las conexiones que ellos describen tener están más bien restringidas a familia extendida y colegas del trabajo con quienes se reúnen principalmente para celebraciones religiosas. Este tipo de redes sociales contrasta con el tipo de comunidad y conexiones observados en la comunidad latina mayoritaria. Esta comunidad en Durham aunque aún descentralizada cuenta con la presencia de organizaciones que les ofrece un punto de asociación en español<sup>5</sup>.

---

<sup>5</sup> Nos referimos a organizaciones como El Centro Hispano, Iglesias católicas o de otras denominaciones, Cooperativas Crediticias, etc.

Este contraste entre tipos de comunidades que diferencian a los grupos originarios de los grupos de habla hispana refleja una extensión de la situación de inequidad y vulnerabilidad vivida en el propio México. Además del factor racial y socioeconómico el factor lingüístico tiene un peso importante para determinar situaciones de poder y de marginación. El creciente bilingüismo en español en México, en detrimento de la mantención de las lenguas originarias, no solo refleja el poco valor social de éstas, sino que además apunta a un incremento de, por un lado, el uso de una variedad de español bilingüe y por otro, la existencia de una variedad monolingüe hablada por las generaciones posteriores a las de los bilingües. Ambas variedades son estigmatizadas por las poblaciones monolingües sin contacto con lenguas originarias (Zimmerman 1999).

### 3. ENTREVISTAS

Las entrevistas realizadas corroboran las percepciones y actitudes respecto de los grupos originarios desde una perspectiva hispanohablante. Uno de los entrevistados hispanohablantes<sup>6</sup> indicó que [en inglés]: “Hay lugares en México donde no hablan español bien... porque no se comunican con la gente, ellos están muy separados... separados”<sup>7</sup> y cuando se le preguntó cómo eran las relaciones entre la gente que solo habla español y la gente que habla una lengua originaria en Durham; su respuesta fue: “No hay comunicación. Nada de eso”<sup>8</sup>.

Desde el punto de vista de selección de lengua de uso cabe resaltar que los hablantes de origen monolingüe en español responden a las preguntas de las entrevistas en inglés, aun cuando éstas han sido planteadas en español. Aunque no exploramos sistemáticamente este hecho, hay dos posibles explicaciones para su ocurrencia. Una, las personas que entrevistaban tenían como lengua materna el inglés por lo que el informante optaba usar la lengua de su interlocutor. Esta explicación asume que el informante utiliza la estrategia de la convergencia en el proceso de acomodación a su interlocutor (Giles, Coupland, y Coupland 1991). De esta manera, tal

---

<sup>6</sup> Este hablante en particular opta por responder en inglés aunque las preguntas fueron planteadas en español.

<sup>7</sup> “There are parts of Mexico, where they don’t speak Spanish well...because they don’t communicate with the other people. They’re very separated...separated”.

<sup>8</sup> “There is no communication, nothing like that”.

estrategia le permite al informante asumir una posición de igualdad de estatus con el interlocutor al hacer uso del inglés, estableciendo además una mayor distancia entre él o ella y el grupo de hispanos hablantes de lenguas originarias.

### 3.1. Lengua e identidad

La situación de inmigración en Estados Unidos parece exacerbar esta diferenciación entre grupos mexicanos, contexto en el cual muchos de la comunidades originarias redefinen su identidad al expresar oposición a lo latino. “No somos *latino*” comentaba uno de los propietarios otomí de negocios en Durham. “Compramos en tiendas diferentes, comemos comida diferente y hablamos una lengua diferente. Ellos [refiriéndose a los mexicanos en Durham] no nos consideran latino, y nosotros no nos consideramos *latino*.” Tal como está implicado en esta declaración, no solo no hay un reconocimiento de los grupos originarios por parte de los latinos, sino que esta ausencia trasciende la esfera social y forma parte de la esfera pública oficial en el año 2009<sup>9</sup>. Examinando un poco más este rechazo de lo latino parece estar asociado al sentimiento de aislamiento que han experimentado en Durham, que para algunos ha sido una extensión de la discriminación sufrida en México. No pertenecen a la categoría *latino* porque su lengua primaria no es el español, lo cual los distingue claramente de los inmigrantes hispanohablantes. Es en esta situación de posicionamiento socio-político donde el poblador originario encuentra la oportunidad de desligarse de su procedencia mexicana.

En este extracto se observa también que el hablante utiliza la versión singular del término latino. Lo que las transcripciones reflejan es la ausencia de la forma plural en contextos requeridos. La ausencia de concordancia de número o género, sin embargo no es extendida en este hablante. Por eso asumimos que el hablante usa el término *latino* en inglés, sin marca de pluralidad. No somos *latino*.

El siguiente ejemplo alude a la falta de reconocimiento de los grupos indígenas en Durham. El hablante es un adulto mexicano con raíces otomíes, pero cuya generación familiar es solo hispanohablante.

<sup>9</sup> En los últimos años el ámbito público ha ido cobrando otro aspecto. Al menos, instituciones como El Centro Hispano han hecho explícita la necesidad de reconocer a las poblaciones que hablan lenguas originarias y han instaurado algún tipo de servicio orientado a necesidades que se creen específicas. Por otro lado, algunos hablantes de lenguas originarias, como el otomí, han entrado en una esfera más pública (aunque académica) a partir de intereses concretos tales como la medicina natural.

- Pregunta: ¿Cuál crees que es la perspectiva general de inmigrantes mexicanos en los Estados Unidos respecto de los hablantes de lenguas indígenas? ¿Hay un estigma vinculado al habla de estas lenguas?
- Respuesta: [en inglés] No hay gente que las hable. Esto fue en la última generación. **Se acabó** [el énfasis es nuestro]. En el campo porque no hay escuelas, mi abuela no sabía leer ni escribir, no había escuela. Pero mi abuela sabía hablar español muy bien- ellos empezaron a civilizarse; la gente empezó a usar las dos lenguas; otomí y español; pero ni mi mamá ni mis tías hablaban otomí. Ellas perdieron el otomí<sup>10</sup>.

Se observa en la respuesta de este hablante, la conexión entre lengua, educación, *civilización* y el supuesto erróneo de que no hay gente que hable estas lenguas. Si bien es posible que esta referencia ambiguamente esté apuntando a una experiencia personal, el resto de la entrevista y su discurso expresa una negación de la realidad sociolingüística indígena. El uso del término *civilización* parece aludir al proceso de ‘modernización’ que está claramente asociado con un proceso de aculturación hacia la cultura dominante hispana en México.

De la misma manera, otra informante hispanohablante entrevistada habla de los grupos originarios en Durham enfatizando la falta de escolaridad como una causa de su situación de desprestigio y marginación. De manera interesante si bien esta persona no se compromete con una opinión fuerte sobre el uso lingüístico de los hablantes de lenguas originarias, utiliza los términos ‘conocimiento’ y ‘cultura’ asociados al nivel de educación.

- Pregunta: ¿Crees que hay dificultades de comunicarse con una persona que habla una lengua indígena?
- Respuesta: [en inglés] Sí, definitivamente. Además de la lengua está también, está el conocimiento. El *conocimiento* de todo , por *la cultura*, porque hay personas que inclusive, que conozco, no de México, pero otros hispanos por ejemplo, y he visto que su cultura no es muy amplia, entonces su habla no es muy amplia [...] mientras mas pequeño tu conocimiento, entonces más tu lengua va a ser pequeño<sup>11</sup>.

<sup>10</sup> “There aren’t people who speak them [lenguas indígenas], this was the last generation —it ended. In the countryside, because there aren’t schools, my grandma didn’t read or write, there wasn’t school. But my grandma spoke Spanish very well— they began to civilize, people then started to speak the two languages —Otomí and Español— but neither my mom nor my aunts spoke Otomi —they lost the Otomi”.

<sup>11</sup> “Yes, definitely, apart from the language, too, it’s the knowledge... Of everything – because of the culture, because there are even people that I know, not from Mexico, but other Hispanics from El Salvador for example,

Así mismo, la perspectiva presentada por la misma persona más adelante en la entrevista alude a la discriminación de estos hablantes por motivos lingüísticos.

[en inglés]... y allí está siempre la discriminación. Estoy segura porque tú la ves en la actitud, etc. ... Tengo, no amigos, pero gente que conozco, conocidos del trabajo, de El Salvador, que hablan raro... y que... y siempre viene la burla. Entonces, imagina que con las personas indígenas es peor, y es triste porque siempre es la burla, con los indios...<sup>12</sup>

### 3.2. Tensiones lingüísticas

En este contexto, entonces, no es difícil entender que la tensión sociolingüística se manifieste también en la percepción de identidad. Hay una clara intención de definir la identidad indígena por disociación del grupo mayoritario.

Otra muestra de esta actitud la sugiere el hecho de que algunos de los entrevistados hablantes de Náhuatl indicaron que no sentían la posibilidad de integrarse a un centro hispano cuando las hojas de registro no contemplaban la alternativa de que los individuos de procedencia mexicana hablaran una lengua originaria como parte de su repertorio lingüístico primario.

Uno de los entrevistados en particular, que reportó tener un nivel de educación bajo, indicó que la relación entre personas que hablan lenguas originarias y sus paisanos monolingües, hispanohablantes cambiaba de alguna manera una vez se cruzaban las fronteras. Con la excepción de lenguas extranjeras, esta persona indicaba que en México el marco lingüístico era bastante simple: español y las lenguas originarias. De acuerdo a este hablante, en México cuando un monolingüe de español escucha una lengua que no conoce éste asume que se trata de una lengua originaria. Con ello se le asigna al hablante de dicha lengua la carga estereotípica que reciben estos pueblos, particularmente en zonas urbanas. Sin embargo, en los Estados Unidos, y particularmente en Durham una ciudad caracterizada por una presencia

---

and I've seen that sometimes their culture isn't very 'amplia'/extensive, then because of that their speech isn't very extensive [...] because then 'mientras mas pequeño tu conocimiento, entonces más tu lengua va a ser pequeño'".

<sup>12</sup> "...there's always the discrimination. I'm sure, because, you see it in the attitude, etc...I have, not friends, but people that I know, acquaintances, that like that, I have some friends from work from El Salvador, and they speak differently, 'raro'...and that...it always comes to mocking. Then, imagine that with an indigenous person it's worse, and it's sad, because it's always 'la burla,' with 'indios'".

significativa de una población internacional (dadas las universidades y los centros de investigación que radican en esta región), los hablantes monolingües de español están expuestos a una diversidad lingüística más amplia. En este contexto, 'el otro' ya no es necesariamente el hablante de lengua originaria sino todo aquel que no habla español; pero a la misma vez, la población hispanohablante se convierte en 'el otro' en una sociedad con un fuerte componente anglo que identifica a esta comunidad por el uso lingüístico del español.

De esta manera, los hablantes de lenguas originarias son trasladados a un doble nivel de 'otredad'. Se intensifica las diferencias con los monolingües mexicanos, y se les añade el sello del inmigrante latino desde la perspectiva anglo. Así, el hablante de lengua originaria es el 'otro' que llega desde México para los hablantes monolingües Mexicanos; y a la vez, es el 'otro' que se incorpora a las masas indiferenciadas de población hispanohablante inmigrante llamada por los estadounidenses, latinos. El sentimiento respecto a una división motivada lingüísticamente es compartido por los hablantes de lenguas originarias que pudimos entrevistar. El siguiente es un extracto de una entrevista con un hablante de otomí.

- Pregunta: ¿siente usted una división de la comunidad latina aquí en Durham por líneas... de diferentes lenguas? ¿Es que hay grupos? ¿Los que hablan otomí están juntos, y los que hablan sólo español están juntos y separados también?
- M.E. Ajá, sí.
- Pregunta: ¿Hay una división?
- M.E. Sí.
- Pregunta: Por eso, ¿la mayoría de sus amigos hablan otomí?
- M.E. Sí, hablan otomí.
- Pregunta: Entonces, cuando habla con sus amigos, ¿habla en otomí?
- M.E. Ajá.
- K: ¿Siempre?
- M.E.: Sí.

La declaración del hablante de otomí, aunque con pocas palabras, confirma la percepción de una división en la comunidad latina, que se traduce a una percepción de estigma que como indicaba uno de los hablantes se trae desde México.

Por otro lado, la selección de lengua de uso diario es importante también porque reafirma uno de los componentes más fuertes de su identidad.

- Pregunta: ¿Con cuál idioma te comunicas más, o qué idioma usas más cada día?
- M: Pues, el otomí, el español, y el inglés no más con los patrones, amigos, así este americanos negros. No más que necesito, pues hablo poquito inglés, no mucho.
- Pregunta: No mucho.
- M: Poquito.
- Pregunta: ¿Cuáles idiomas habla su esposa?
- M: Otomí y español, nada más.
- Pregunta: ¿Ella no habla inglés?
- M: Ella apenas puede decir “sí” o “no” en inglés pero...
- Pregunta: ¿Ella habla muy bien el español también?
- M: No muy bien el español.

### 3.3. Lengua originaria y uso lingüístico

El uso lingüístico varía según edad del hablante, nivel de educación, relación con el interlocutor y el dominio que tenga de la lengua originaria. En general, mientras más joven y más educado o educada la persona, el uso de esta lengua será más reducido o en dominios más restringidos. Algunos de los hablantes indicaron que la vergüenza lingüística afectaba negativamente el uso de las lenguas originarias en las generaciones más jóvenes. El siguiente cuadro resumen la información sobre selección lingüística:

Hablante	Lengua dominante (reportada)	Lengua usada con el esposo o esposa	Lengua usada con los hijos	Lengua usada en el trabajo	Lengua usada con amigos
M.B.	otomí	otomí	otomí/español	otomí/inglés	otomí
M.E.B.	otomí	otomí	otomí/español	español/inglés	otomí
R.A.	otomí	otomí	n/a	otomí/inglés	otomí
B.H.	español/otomí	n/a	n/a	español /otomí	español
G.E.	otomí	otomí	español	español	otomí
F.E.	náhuatl	español	n/a	español	náhuatl/ español
R.E.	español	español	n/a	español	español
R.C.	chatino/español	n/a	español	español /inglés	chatino

**Cuadro 2.** Selección lingüística.

Es interesante observar que los hablantes que tienen hijos deciden o usar ambas lenguas o solamente español. Sin embargo, en contextos en los que la interacción es con adultos la presencia de la lengua originaria es clara todavía.

#### **4. CONCLUSIÓN**

Los resultados de estas entrevistas muestran a una población que habla lenguas originarias en Durham dentro de tres marcos lingüísticos que divergen en el valor asignado por los propios hablantes originarios. Como hablantes nativos de las lenguas originarias se identifican éstas en contextos sociales y culturales (es decir, en las relaciones conyugales, con amigos). Esto está también determinado por el grupo de edad al que pertenecen sus contactos. Sin embargo, como residentes en una comunidad anglohablante, uno de los objetivos de estos hablantes de lenguas originarias es lograr interactuar en inglés con la población de Durham. Este objetivo es relevante más allá de su proficiencia lingüística actual en esta lengua.

El segundo marco de uso lingüístico, el uso de español - puede ser el más complicado de los tres. Ya que no es su lengua materna ni la lengua mayoritaria en Durham (o de prestigio), español pareciera existir como un inconveniente lingüístico para la reafirmación y validación de una identidad que tratan de mantener y rescatar en este contexto multilingüe. El español es necesario para la comunicación con la comunidad latina mayoritaria; por ejemplo funciona como una segunda lengua de enseñanza en la escuela o en los servicios de salud en los hospitales o clínicas. Más aún, el español tiene un tercer propósito en el contexto donde inglés no es accesible para estos hablantes, como lengua de comunicación con la población anglo educada de Durham, mucha de la cual habla esta lengua o tienen acceso a recursos en español. Esta faceta particular del panorama sociolingüístico de los hablantes de lenguas originaria reitera y refuerza la situación de opresión lingüística que se reproduce en este país para los hablantes de lenguas originarias.

La situación sociolingüística de los inmigrantes que hablan lenguas originarias en Durham permite examinar procesos de aculturación desde una perspectiva dinámica y fluida que involucra a todos los elementos que interactúan en la sociedad —lingüísticamente o no. A la misma vez, estos procesos de aculturación y dicha interacción sirven de trasfondo para que un grupo de individuos negocien y redefinan su identidad.



Lo que se ha observado en el caso de los hablantes originarios es que descubren que son doblemente estigmatizados en la sociedad de Durham. El efecto que esto tiene en su definición de identidad es el rechazo a ser considerado latino y la oportunidad de reafirmar su origen indígena, al menos entre los hablantes de la misma generación. Esto está determinado por su selección lingüística.

Sin embargo, esta confirmación y reafirmación de la identidad indígena no exime la situación de discriminación y la reiteración de inequidades sufridas históricamente<sup>13</sup>. La falta de acceso a servicios a menos que se use español es un ejemplo de ello.

Por último, este trabajo presenta una situación en la que un grupo marginal inmigrante negocia su identidad como parte del proceso de integración en una comunidad estadounidense. Queda por decir que el aceptar la presencia sociolingüística y cultural de los hablantes de lenguas originarias y, por lo tanto darle un reconocimiento oficial no desafía las fronteras lingüísticas tradicionales si éstas continúan reproduciendo inequidades.

---

<sup>13</sup> Tal como proponen Padilla y Pérez (2003) individuos que son claramente identificables como foráneos (lingüísticamente y racialmente) son más vulnerables a ser objeto de prejuicio y discriminación por los grupos dominantes.

## REFERENCES

- Cruz, Yolanda (2006): “Sueños binacionales”, en L. Burgueno (prod.): *All Roads Film Festival Collection*. DVD. Washington, DC: National Geographic.
- Giles, Howard, Justine Coupland y Nikolas Coupland. (1991): *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Cambridge: Cambridge University Press.  
<http://dx.doi.org/10.1017/CBO9780511663673>
- Padilla, Amado M. y William Pérez (2003): “Acculturation, social identity, and social cognition: A new perspective”. *Hispanic Journal of Behavioral Sciences* 25, pp. 35-55.  
<http://dx.doi.org/10.1177/0739986303251694>
- Zimmerman, Klaus. (1999): *Política del lenguaje y planificación para los pueblos amerindios*. Madrid: Iberoamericana.



