LINKING COMPREHENSION COSTS TO PRODUCTION PATTERNS DURING THE PROCESSING OF MIXED LANGUAGE

A Dissertation in
Spanish and Language Science
by
Rosa E. Guzzardo Tamargo

© 2012 Rosa E. Guzzardo Tamargo

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

August 2012
The dissertation of Rosa E. Guzzardo Tamargo was reviewed and approved* by the following:

Paola E. Dussias
Associate Professor of Spanish, Linguistics, and Psychology
Dissertation Co-Advisor
Co-Chair of Committee

Rena Torres Cacoullos
Professor of Spanish and Linguistics

John M. Lipski
Edwin Erle Sparks Professor of Spanish and Linguistics

Judith F. Kroll
Distinguished Professor of Psychology, Linguistics, and Women’s Studies

Chip Gerfen
Associate Professor of Spanish and Linguistics
Head of the Department of Spanish, Italian and Portuguese
Dissertation Co-Advisor
Co-Chair of Committee

*Signatures are on file in the Graduate School
ABSTRACT

Studies on codeswitching have focused mainly on production. The experiments reported in this dissertation focus, instead, on how Spanish-English bilinguals process codeswitches during reading comprehension. Two types of intrasentential codeswitches are examined here: a switch between the Spanish progressive auxiliary *estar* ‘be’ and an English present participle versus a switch between the Spanish perfect auxiliary *haber* ‘have’ and an English past participle. Despite their superficial similarities, *estar*+English participle switches are more frequent in natural bilingual production corpora than *haber*+English participle switches. The main objective of this dissertation is to examine whether comprehension costs reflect production patterns, by investigating if the more frequent switches involving the progressive structure are easier to process than the less frequent switches involving the perfect structure. This set of studies also examines whether age of second language acquisition and task play a role in the way these two types of codeswitches are processed.

Groups of early bilinguals and late bilinguals took part in an eye-tracking study in which they were asked to read sentences to complete two tasks: a comprehension task and an acceptability judgment task. The participant groups exhibited similar results in the comprehension task. That is, *estar*+English participle switches were processed more easily than *haber*+English participle switches, reflecting the tendencies found in natural production. In the acceptability judgment task, the early bilinguals no longer displayed sensitivity to the distributional patterns of these different types of switches, and instead processed both of them with similar difficulty. The late bilinguals, however, were less affected by the task and, therefore, processed the codeswitches in the acceptability judgment task similarly to the way that they did in the comprehension task. Both early and late bilinguals displayed no switch costs when reading codeswitched sentences, compared to unilingual sentences. Overall, these results lend support for
constraint satisfaction models of sentence processing (MacDonald & Seidenberg, 2006), such as the Production-Distribution-Comprehension approach (Gennari & MacDonald, 2009; MacDonald & Thornton, 2009). When bilinguals complete a task that resembles natural comprehension, the comprehension costs that they exhibit are directly related to the distributional patterns of the codeswitches in natural production.
TABLE OF CONTENTS

LIST OF FIGURES .................................................................................................................. viii

LIST OF TABLES .................................................................................................................. ix

ACKNOWLEDGMENTS ......................................................................................................... xi

Chapter 1 Codeswitching: From grammatical studies on production to psycholinguistic studies on comprehension ................................................................. 1

1.1 The grammatical study of codeswitching ................................................................. 2
   1.1.1 Linear constraints on codeswitching .............................................................. 3
   1.1.2 Hierarchical constraints on codeswitching .................................................... 5

1.2 The psycholinguistic study of codeswitching ......................................................... 7

1.3 Codeswitches involving the auxiliary phrase ....................................................... 10
   1.3.1 Unacceptability of auxiliary phrase switches ............................................... 10
   1.3.2 Acceptability of auxiliary phrase switches .................................................. 11

1.4 Asymmetry in naturalistic production of codeswitches ......................................... 12
   1.4.1 Asymmetry in naturalistic production of auxiliary phrase switches ............ 13
   1.4.2 Accounting for the asymmetrical production of auxiliary phrase switches .... 14
      1.4.2.1 Morpheme classification account ......................................................... 15
      1.4.2.2 Grammaticalization account ............................................................... 16
      1.4.2.3 Structural distance account ................................................................. 20

1.5 Comprehension of codeswitches ............................................................................ 23

1.6 Relating comprehension costs and production patterns ...................................... 25

1.7 Justifying a focus on reading comprehension ....................................................... 28

1.8 Relating reading and auditory sentence comprehension ..................................... 30

1.9 Task effects on sentence processing ..................................................................... 31

1.10 Aims of this dissertation ....................................................................................... 36

Chapter 2 Spanish-English auxiliary phrase switches in oral and written corpora ......... 38

2.1 Data ............................................................................................................................. 39
   2.1.1 Oral corpus ..................................................................................................... 39
   2.1.2 Written corpus ............................................................................................... 40
   2.1.3 Differences between the oral and written corpora ......................................... 41

2.2 Extraction and coding methods ............................................................................. 44

2.3 Results ......................................................................................................................... 46
   2.3.1 Oral corpus ..................................................................................................... 46
   2.3.2 Written corpus ............................................................................................... 49

2.4 Discussion .................................................................................................................... 51
Chapter 3 Comprehension costs associated with auxiliary phrase switches

3.1 Research questions ........................................................................................................56
3.2 Method ...........................................................................................................................57
   3.2.1 Participants .............................................................................................................57
      3.2.1.1 Language History Questionnaire .................................................................58
      3.2.1.2 Boston Naming Vocabulary Test .................................................................58
      3.2.1.3 Grammatical competency tests .................................................................59
   3.2.2 Materials and Design ...........................................................................................60
   3.2.3 Procedure ..............................................................................................................62
3.3 Predictions .....................................................................................................................63
3.4 Results ..........................................................................................................................66
   3.4.1 Results for the comprehension task .................................................................68
      3.4.1.1 Early bilinguals .........................................................................................69
      3.4.1.2 Late bilinguals .........................................................................................71
   3.4.2 Results for the acceptability judgment task ..................................................73
      3.4.2.1 Early bilinguals .........................................................................................75
      3.4.2.2 Late bilinguals .........................................................................................75
3.5 Discussion ......................................................................................................................76
3.6 Conclusion ......................................................................................................................82

Chapter 4 Comprehension costs associated with auxiliary phrase switches in an established community of codeswitchers

4.1 Research questions ........................................................................................................85
4.2 Method ..........................................................................................................................86
   4.2.1 Participants ...........................................................................................................86
   4.2.2 Materials and Design .........................................................................................90
   4.2.3 Procedure .............................................................................................................93
4.3 Predictions .....................................................................................................................93
4.4 Results ..........................................................................................................................94
   4.4.1 Results for the comprehension task .................................................................94
      4.4.1.1 Early bilinguals .........................................................................................95
      4.4.1.2 Late bilinguals .........................................................................................96
   4.4.2 Results for the acceptability judgment task ..................................................98
      4.4.2.1 Early bilinguals .........................................................................................99
      4.4.2.2 Late bilinguals .........................................................................................100
   4.4.3 Reading speed results .........................................................................................102
4.5 Discussion .....................................................................................................................105
4.6 Conclusion .....................................................................................................................111
Chapter 5  Switch costs............................................................................................................. 113
  5.1 Method............................................................................................................................. 120
    5.1.1 Participants.................................................................................................................. 120
    5.1.2 Materials and Design............................................................................................... 121
    5.1.3 Procedure.................................................................................................................. 123
  5.2 Results.............................................................................................................................. 123
    5.2.1 Results for the comprehension task........................................................................... 124
      5.2.1.1 Early bilinguals................................................................................................. 124
      5.2.1.2 Late bilinguals................................................................................................. 125
    5.2.2 Results for the acceptability judgment task.............................................................. 127
      5.2.2.1 Early bilinguals................................................................................................. 127
      5.2.2.2 Late bilinguals................................................................................................. 128
  5.3 Discussion......................................................................................................................... 129

Chapter 6  Conclusions and future directions ........................................................................ 134

References.................................................................................................................................. 146

Appendix A  Sample questions from the Language History Questionnaire ....................... 159

Appendix B  Examples of the outline drawings included in the Boston Naming
  Vocabulary Test...................................................................................................................... 160

Appendix C  Sample items from the grammatical competency tests.................................... 161

Appendix D  Codeswitched experimental item sets ................................................................. 164

Appendix E  Unilingual experimental item sets ..................................................................... 174
LIST OF FIGURES

Figure 1-1: Syntactic representation of the Spanish progressive structure. ..............................21

Figure 1-2: Syntactic representation of the Spanish perfect structure. .................................22

Figure 6-1: Mean reading times at the participle by condition for the State College (SC) participants during the comprehension task. .................................................................136

Figure 6-2: Mean reading times at the participle by condition for the New York (NY) participants during the comprehension task. .................................................................136

Figure 6-3: Mean reading times at the auxiliary by condition for the New York (NY) participants during the comprehension task. .................................................................139

Figure 6-4: Mean reading times at the participle by condition for the early bilinguals in State College (SC) and New York (NY) during the acceptability judgment task. .........141

Figure 6-5: Mean reading times at the participle by condition for the late bilinguals in State College (SC) and New York (NY) during the acceptability judgment task. .........142
LIST OF TABLES

Table 1-1: Examples of *estar*+English participle and *haber*+English participle switches. .....23

Table 1-2: A comparison between written and oral codeswitch types.................................29

Table 2-1: Distribution in the oral corpus of codeswitches in sentences including the progressive structure. .................................................................46

Table 2-2: Distribution in the oral corpus of codeswitches in sentences including the perfect structure..................................................................................48

Table 2-3: Distribution in the written corpus of codeswitches in sentences including the progressive structure. .................................................................49

Table 2-4: Distribution in the written corpus of codeswitches in sentences including the perfect structure..........................51

Table 3-1: Characteristics of the State College participant groups........................................60

Table 3-2: Example of experimental item set. ........................................................................62

Table 3-3: Proportion of correct answers to the comprehension questions by condition and participant group. .................................................................68

Table 3-4: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the early bilinguals during the comprehension task. .....................69

Table 3-5: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the late bilinguals during the comprehension task. .......................71

Table 3-6: Proportion of acceptable judgments of the sentences by condition and participant group. ............................................................................................72

Table 3-7: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the early bilinguals during the acceptability judgment task. ..............73

Table 3-8: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the late bilinguals during the acceptability judgment task. ....................75

Table 3-9: Summary of the State College eye-tracking results by reading measure, participant group, and task.................................................................77

Table 4-1: Characteristics of the New York participant groups. .............................................89
Table 4-2: Examples of unilingual English sentences (matched translations of the
codeswitched sentences). ................................................................. 90

Table 4-3: Proportion of correct answers to the comprehension questions by condition
and participant group. ........................................................................... 94

Table 4-4: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the early bilinguals during the comprehension task. ......................... 95

Table 4-5: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the late bilinguals during the comprehension task. .......................... 97

Table 4-6: Proportion of acceptable judgments of the codeswitched sentences by
condition and participant group. ..................................................................... 98

Table 4-7: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the early bilinguals during the acceptability judgment task. .............. 99

Table 4-8: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the late bilinguals during the acceptability judgment task. ............... 100

Table 4-9: Mean total reading time (in milliseconds) of the experimental sentences by
structure/language and participant group for the comprehension task.................... 102

Table 4-10: Mean total reading time (in milliseconds) of the filler sentences by language
and participant group for the comprehension task. ............................................ 103

Table 4-11: Mean total reading time (in milliseconds) of the experimental sentences by
structure/language and participant group for the acceptability judgment task........... 104

Table 4-12: Mean total reading time (in milliseconds) of the filler sentences by language
and participant group for the acceptability judgment task. ................................... 104

Table 4-13: Summary of the New York eye-tracking results by reading measure,
participant group, and task. ............................................................................. 106

Table 5-1: Sample sentences used to examine switch costs. ................................... 121

Table 5-2: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the early bilinguals during the comprehension task. ......................... 124

Table 5-3: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the late bilinguals during the comprehension task. .......................... 126

Table 5-4: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the early bilinguals during the acceptability judgment task. .............. 127

Table 5-5: Mean gaze duration, regression path time, and total time (in milliseconds) by
condition for the late bilinguals during the acceptability judgment task. ............... 128
ACKNOWLEDGMENTS

I greatly appreciate the time and effort that the members of my dissertation committee contributed to this work: Paola E. Dussias, Chip Gerfen, Rena Torres Cacoullos, John M. Lipski, and Judith F. Kroll. Each committee member had a necessary role in developing this dissertation and their guidance was important for its successful completion. My main advisors, Paola E. Dussias and Chip Gerfen, have been my mentors during my entire time in the Graduate Hispanic Linguistics Program at Penn State University. Their support and encouragement have always sustained me to complete my degree and to continue to grow professionally. I would also like to express my gratitude to several members of The City College of New York whose help was instrumental for my data collection in New York: Pamela Gillespie (Associate Dean and Chief Librarian), Ebe P. De Leon (CUNY Office Assistant), Simon Simms (Chemistry Department), Laura Callahan (Department of Foreign Languages and Literatures), and Lissy Wassaf (CUNY UI-IRB Administrator).

Furthermore, I would like to acknowledge several individuals with whom I consulted throughout the design and development of this set of studies as well as the analyses of results: Maya Misra, Jason Gullifer, Kosta Dussias, Rosanna Otondo, and Eleonora Rossi. Special thanks are also due to the Penn State undergraduate research assistants who helped me to collect and extract data, particularly Marie-Christine Thebèrge and Victoria Rodríguez, as well as to all the participants who took part in this set of studies. Moreover, I would like to thank my family, who has provided a great deal of reassurance as I have advanced professionally. I am grateful to my fiancé, Javier, my parents, Maribel and John, and my sisters, Mariana, Paloma, and Isabel. Their words of encouragement and their unconditional help were invaluable to complete this work; they
truly helped me reach my goals. I also extend my appreciation to my aunt and uncle, Conchi and Mac, who helped me carry out the idea of collecting data in New York. I look back on the time I spent at their house with very fond memories. Finally, I want to note that this research was achieved thanks to funding provided by the National Science Foundation (NSF Grant BCS-0821924 to Paola E. Dussias and Chip Gerfen, NSF Dissertation Improvement Award BCS-1123874), and by the College of the Liberal Arts and the Center for Language Science at Penn State University.
Chapter 1

Codeswitching: From grammatical studies on production to psycholinguistic studies on comprehension

One of the outcomes of contact between bilinguals’ two languages is the act of codeswitching, in which they spontaneously alternate between both linguistic systems. Some initial reports viewed this phenomenon as a sign of incomplete acquisition of the two languages, which resulted in mental confusion and inability to separate both systems. Codeswitching was considered part of the performance of a nonfluent bilingual and an indication of laziness and lack of education. For instance, in his study on Spanish-English bilingualism in Texas, Lance (1969) cited a magazine article of that time (“Nation: The little strike that grew to La Causa,” 1969) and referred to the common unjustified assertion that “Mexican-Americans in Texas speak neither English nor Spanish but, instead, a random, grammarless mixture of the two, pejoratively referred to as ‘Tex-Mex’; an alleged ‘frequent result’ is ‘that they become not bilingual but nearly nonlingual’.” However, soon after, more systematic studies found that codeswitching was not haphazard interference between both languages, but rather a natural, smooth, and fluent phenomenon that displayed controlled integration of both linguistic systems (Lance, 1969, p. 93; Lipski, 1982, p. 191, 2005, p. 17). As a result, codeswitching has become an increasingly popular topic of study and it has been examined from many different perspectives. Studies that take a sociolinguistic viewpoint are principally concerned with the social, stylistic, and pragmatic functions of codeswitching (e.g., Blom & Gumperz, 1982; McClure & McClure, 1988; Myers-Scotton, 1993a; Valdés-Fallis, 1976; Zentella, 1997). Others are geared towards developing typologies of codeswitching (e.g., Clyne, 2003; Muysken, 2000). Yet another group of studies, which focuses on the grammatical aspects of codeswitching, aims to describe and quantify the
elements in an utterance that can be codeswitched (e.g., Belazi, Rubin, & Toribio, 1994; Di Sciullo, Muysken, & Singh, 1986; Lipski, 1978; MacSwan, 2000; Pfaff, 1979; Poplack, 1980; Sankoff & Poplack, 1981; Timm, 1975). Finally, studies that take a psycholinguistic standpoint strive to explain codeswitching within existing language processing models (e.g., Azuma, 1993; Joshi, 1985; Myers-Scotton, 1993b, 2006). Several studies belonging to these last two types will be further described below.

A fundamental distinction in the codeswitching literature takes place between intersentential switches and intrasentential switches. In the first case, language switching occurs at sentence boundaries (e.g., I need some fresh air. Voy a pasear en el parque. ‘I’m going to take a walk in the park.’). In the second case, the alternation from one language to the other occurs within sentence boundaries (e.g., Mi hermano me dijo que he wouldn’t make it home in time for dinner. ‘My brother told me that he wouldn’t make it home in time for dinner.’). Since the latter type of switches requires greater simultaneous control of both languages and displays stronger interaction between both linguistic systems, it has provided an ideal testing ground for putative constraints on codeswitching. This dissertation focuses precisely on this type of codeswitches, specifically on switches between auxiliaries and participles. The following sections will describe different studies that have taken grammatical and psycholinguistic approaches to study codeswitching.

1.1 The grammatical study of codeswitching

As mentioned above, many researchers have focused on determining the linguistic principles that guide codeswitching (e.g., Belazi, Rubin, & Toribio, 1994; Di Sciullo, Muysken, & Singh, 1986; Lipski, 1978; MacSwan, 2000; Pfaff, 1979; Poplack, 1980; Sankoff & Poplack, 1981; Timm, 1975).

1 Throughout this dissertation, Spanish words and phrases are shown in italics.
1981; Timm, 1975). The data used in these studies consists mainly of recordings of natural and informal conversations or interviews and of grammaticality or acceptability judgments, as their main goal is to describe points in an utterance at which codeswitching does and does not occur, in order to define constraints on codeswitching.

1.1.1 Linear constraints on codeswitching

Early codeswitching studies postulated linear constraints on codeswitching that were based on the linear surface order of the sentence constituents. Timm’s (1975) study on Spanish-English bilinguals living in California constitutes one of the pioneering works to address codeswitching not as random and unsystematic behavior, but as a phenomenon that obeys a stringent set of extralinguistic and linguistic rules. The researcher analyzed the switch sites in bilingual oral exchanges and written texts. She also elicited acceptability judgments from bilinguals for test sentences with switches at different syntactic sites. Timm found that certain Spanish-English switches did not occur in the bilingual oral and written corpora and were judged unacceptable by Mexican-American speakers when they occurred between pronominal subjects or objects and the verb to which they belonged, between finite verbs and their infinitival complements, between auxiliaries and main verbs, between verbs and negative elements, and in long noun phrases composed of determiners, nouns, and adjectives. These observations contributed a set of testable descriptive generalizations for future codeswitching studies.

In the late 1970s, several researchers independently proposed a linear constraint on codeswitching, which inhibits switching from occurring in contexts with word order differences between the languages in contact. Lipski (1978) analyzed examples of Spanish-English codeswitched utterances, collected either from personal observation or from the codeswitching studies of earlier researchers. Meticulous examination of these sentences, coupled with
acceptability judgments gathered from native speakers, led Lipski to propose a necessary condition for bilingual codeswitching: superficial homology or syntactic congruence in the portion of both phrases that falls after the actual switch (p. 254). In other words, although the portion of a mixed utterance that falls before the codeswitch can be syntactically divergent in both languages, those portions that fall after the switch must have syntactically identical surface structures. For example, the following codeswitched sentence is acceptable because, although the initial portion of the sentence is syntactically different in English and Spanish (i.e., the adjective precedes the noun in English, but follows it in Spanish), the portion after the switch is syntactically identical in both languages.

(1) Codeswitched sentence: *Mi hermano mayor* went to the store on the corner.
   English sentence: *My older brother* went to the store on the corner.
   Spanish sentence: *Mi hermano mayor fue a la tienda de la esquina.*

In trying to address the shortcomings of previous codeswitching studies, namely limited conversation samples and unreliable acceptability judgments, Pfaff (1979) examined conversational data from Spanish-English bilingual speakers of various ages and social backgrounds from California and Texas. For purposes of comparison, Pfaff also analyzed two examples of written data. Based on a quantitative analysis of relative frequencies of switches at different syntactic junctures, Pfaff proposed four types of constraints on codeswitching: functional constraints, structural constraints, semantic constraints, and discourse constraints. For instance, codeswitches can occur as long as they do not interfere with the expression of tense/aspect/mood and subject/object relations. In addition, codeswitching is more common at syntactic sites where the superficial linear structure of the two languages coincides. Moreover, non-clausal switches tend to start at a point of lexical divergence (e.g., before main verbs, nouns, or adjectives). Finally, certain less common types of switches (e.g., noun phrase switches including the determiner and the noun) are favored in cases of first mention or emphasized mention.
As part of a larger sociolinguistic project involving ethnographic, attitudinal, and linguistic observation, Poplack (1980) conducted a multivariate statistical analysis of the oral codeswitching behavior of Puerto Rican speakers of varying degrees of bilingual ability, who lived in a stable Puerto Rican bilingual community in East Harlem, New York. The results showed that both balanced and unbalanced bilinguals adhered to the Free Morpheme Constraint, which states that a codeswitch can occur after any constituent in discourse as long as the constituent is not a bound morpheme. Both groups of bilinguals also followed the Equivalence Constraint, which states that codeswitches tend to occur at points in discourse where juxtaposition of the elements in both languages does not violate a syntactic rule of either language. Moreover, Poplack found that codeswitching could be used as an indicator of bilingual ability. While the Spanish-dominant bilinguals produced more emblematic switches (i.e., interjections [e.g., shit!], fillers [e.g., I mean], tags [e.g., you know], idiomatic expressions [e.g., no way]), the balanced bilinguals tended to produce more complex intrasentential switches (p. 607).

1.1.2 Hierarchical constraints on codeswitching

The following set of studies sought to determine restrictions on codeswitching based on the structural relations between sentence constituents, as stipulated under Universal Grammar (Chomsky, 1965). Di Sciullo, Muysken, and Singh (1986) explained codeswitching within Chomsky’s syntactic theory of Government and Binding. Following a presentation of several problems with the Equivalence Constraint, the authors proposed that codeswitching is constrained at the surface structure by the government relation that exists between sentence constituents. According to their universally applicable Government Constraint, codeswitching is not possible between elements that hold a government relation. However, whenever the government relation does not hold, lexical elements can be drawn from different lexicons (p. 4). For example,
codeswitching is disallowed between verbs, nouns, or prepositions and their respective complements (particularly the head in the case of complement phrases). In the sentence *I saw the man*, the verb *saw* and the head of its complement phrase, that is, the determiner *the* must be pronounced in the same language. Nonetheless, codeswitching can take place between a noun phrase and a verb phrase, between a determiner and a noun, or between an auxiliary and a verb, among others (p. 12). The data used as support for the Government Constraint consisted of French-Italian-English codeswitched phrases, taken from a naturalistic corpus of speakers living in Montreal, as well as Hindi-English codeswitched phrases, created by the researchers and judged by bilingual informants. The authors explained that the constraint is not the only possible codeswitching restriction, rather that specific languages can impose additional language-particular constraints, which do not override or violate the Government Constraint, but simply complement it.

The study conducted by Belazi, Rubin, and Toribio (1994) constitutes an account for codeswitching under Chomsky’s Minimalist Program. Similar to Di Sciullo, Muysken, and Singh (1986), these authors stated that the constraints on codeswitching should be specified in hierarchical terms and that they should coincide with distinctions and relations already present in the grammar (p. 228). In this work, the authors referred to four previously proposed codeswitching constraints, the Government Constraint, the Free Morpheme Constraint, the Equivalence Constraint, and the Constraint on closed class items (discussed in section 1.2), and they used new data to demonstrate how the constraints were empirically and theoretically inadequate. The data were drawn from two earlier studies, which comprised Tunisian Arabic-French and Spanish-English codeswitched sentences, collected by eliciting grammaticality judgments. Based on these judgmental data, the Functional Head Constraint was put forward. As part of this constraint, a language feature is added to the set of features that must be checked during functional selection. The constraint states that a functional head requires that the language
feature of its complement match its own language feature (p. 228). Therefore, codeswitching is disallowed between functional heads and their complements (e.g., between a complementizer and an inflectional phrase, between a determiner or a quantifier and a noun phrase, between a negative particle or an auxiliary and a verb phrase), but it is allowed between lexical heads (e.g., verbs, prepositions) and their complements.

Based on data gathered from naturalistic and elicited Spanish-Nahuatl corpora, MacSwan (2000) presented a theory of codeswitching that also relies on the concepts and principles of the Minimalist Program. In the Minimalist framework, all morphologically simple and complex items are formed within the lexicon. In addition, the grammatical system is invariant across languages and only allows for parametric variation. According to MacSwan’s approach, the same grammatical relations and operations are relevant to monolingual and bilingual linguistic behavior alike. As a result, nothing constrains codeswitching apart from the requirements of the mixed grammars (p. 44). Since, under this interpretation, all syntactic variation is associated with the lexicon, codeswitching is perceived as the consequence of mixing two languages in the course of a derivation. The requirements of each language are satisfied as the items are selected from a unified lexicon. Because bilinguals have separate phonological systems that cannot be mixed, codeswitching must take place before the utterance reaches the phonological component of the derivation. This approach predicts that codeswitching within heads is not permitted. MacSwan also specifies that language mixture within verb-verb sequences is banned under this account, as well as switches between clitics (both object clitics and negative particles) and verbs (p. 47-48).

1.2 The psycholinguistic study of codeswitching

Instead of focusing on bilingual behavior or bilingual competence, other researchers have examined codeswitching from a processing perspective (e.g., Azuma, 1993; Broersma & De Bot,
These researchers have sought to explain codeswitching in terms of the cognitive processes underlying the production of speech, with the idea that this bilingual act can reveal information about the way different languages are organized in the mind.

Joshi (1985) revisited an idea that had been mentioned by other researchers (e.g., Poplack, 1980, p. 597): the asymmetric relation between the two languages involved in codeswitching. Specifically, Joshi stated that both languages are not equally represented during codeswitching. There is a matrix language, the language that the mixed sentence comes from, and an embedded language, from which some elements are included in the mixed sentence (p. 191). Based on Marathi-English codeswitched sentences, produced judgmentally by the author and two additional bilingual informants, Joshi proposed two general constraints on codeswitching. The first constraint refers to asymmetry during codeswitching: switching a category of the matrix language grammar to the embedded language grammar can occur, but the contrary cannot. The second constraint comprises the nonswitchability of certain categories: open-class items (e.g., nouns, adjectives, verb stems) can be switched, but closed-class items (e.g., determiners, quantifiers, prepositions, tense morphemes, auxiliaries, complementizers, pronouns) cannot (p. 193).

Azuma (1993) further built on Joshi’s proposal, using data from several different language pairs (Swahili-English, Spanish-English, English-Polish, and Hebrew-Spanish) gathered from previous publications as well as naturalistic and judgmental data from Japanese-English bilinguals living in Austin, Texas and Salt Lake City, Utah. Informed by previous speech production models (Bock, 1982; Garrett, 1988; Levelt, 1989), the author put forward the Frame-Content Hypothesis, which states that there are two main stages in production. During the initial planning frame-building stage, closed-class items and morphemes are accessed and retrieved, leaving vacant slots for content words. During the later content-word insertion stage, content
words are introduced into the planning frame. According to Azuma, intrasentential
codeswitching is explained by adopting the Frame-Content Hypothesis and establishing that
codeswitching takes place at the stage of content-word insertion, in which the content words from
a particular language may be inserted into the available slots in the planning frame of another
language (p. 1074). Azuma also suggested that the relation between both stages of production is
strictly serial and noninteractive, that is, there is no feedback from the content-word insertion
stage that could change the frame (p. 1078). As a result, Azuma’s proposal is identical to Joshi’s:
open-class items may participate in codeswitching, but closed-class items cannot.

The Matrix Language Frame (MLF) model and the 4-M model of morpheme
classification (Myers-Scotton, 1993b; Myers-Scotton & Jake, 2000a, 2000b, 2001) are also
informed by psycholinguistic models of language processing that characterize spoken production
(Garrett, 1975; Grosjean, 1988; Levelt, 1989). The MLF model’s structural constraints on
codeswitching are based on two basic oppositions: the Matrix language (ML) versus the
Embedded language (EL), and content morphemes versus system morphemes. The ML is the
more highly activated language, which sets the morphosyntactic frame of the codeswitched
sentence or complementizer phrase, whereas the EL provides singly occurring content elements
or full constituents called EL islands.

The difference between the ML and the EL in codeswitched structures is itself grounded
in a distinction between content and system morphemes. Content morphemes, such as nouns,
most verbs, and some prepositions, are thematic role assigners or receivers that convey semantic
and pragmatic notions. System morphemes, which do not assign or receive thematic roles, are
divided into early and late system morphemes. Early system morphemes (e.g., articles,
derivational affixes) are indirectly conceptually activated and contain structure for conveying the
speaker’s intentions. Late system morphemes contain grammatical information to express
relations between content morphemes. They are further divided into two types. Bridge late
system morphemes unite morphemes into larger constituents, showing their hierarchical relationship. Outsider late system morphemes co-index relations between elements and depend on information outside of their maximal projection in order to integrate morphemes into larger constituents. Typical examples of late system morphemes are certain prepositions, such as English *of*, and inflectional affixes that encode tense. The MLF model and the 4-M model, therefore, generally coincide with earlier proposals (e.g., Azuma, 1993; Joshi, 1985) by taking into account the level of activation of each language during the mental processes involved in language production. Once the matrix language’s morphosyntactic frame is set, elements from the embedded language can be inserted into the pertinent slots.

### 1.3 Codeswitches involving the auxiliary phrase

The present dissertation focuses particularly on codeswitches involving the auxiliary phrase, that is, switches between two Spanish auxiliaries, *estar* ‘be’ and *haber* ‘have’, and their respective English present and past participles. The studies mentioned in the previous section make different predictions regarding the acceptability of these types of codeswitches.

#### 1.3.1 Unacceptability of auxiliary phrase switches

Some of the studies that have relied on grammaticality or acceptability data have categorized auxiliary phrase switches as unacceptable switches. For instance, Timm (1975) found that both *estar*+English participle switches and *haber*+English participle switches were considered anomalous by her bilingual informants (p. 478). Belazi, Rubin, and Toribio (1994) also classified auxiliary+participle switches as unacceptable switches. However, the examples that these researchers provided only include a switch between the perfect auxiliary and the past
participle (pp. 225, 230), and not between the progressive auxiliary and the present participle. MacSwan (2000) used a codeswitched example provided by Belazi et al. (*The students had visto la película italiana. ‘The students had seen the Italian movie.’) to specify that, generally, complex verb-verb sequences are proscribed under his Minimalist approach to codeswitching. However, he did explain that the specific languages involved in codeswitching may change the grammatical status of the switch because language-particular differences are lexically encoded in the Minimalist approach (p. 47). Although, under his account, the switch between the English auxiliary *have* and the Spanish participle is ungrammatical (because English *have* triggers obligatory reanalysis with the Spanish participle, forms a complex $X^0$, which is input to the phonological system, and the derivation crashes because switching is not allowed in the phonological system), the author did not directly refer to switches between Spanish auxiliaries and English participles. It may be that all or some of these auxiliaries behave like the Italian auxiliary *avere* ‘have’, which permits, but does not require reanalysis, therefore, allowing a switch between it and a past participle in another language, such as French (p. 48, e.g., No, parce que hanno donné des cours ‘No, because (they) have given lectures’; in this example, Italian words are shown in normal type and French words are shown in italics).

1.3.2 Acceptability of auxiliary phrase switches

Other studies that include bilingual judgments have, instead, concluded that auxiliary+participle switches are allowed. For example, Di Sciullo, Muysken, and Singh’s (1986) Government Constraint predicts that these switches are possible, provided that one assumes that there is an Inflection node that dominates auxiliaries and that such a node is dominated by the Sentence node (p. 15). If this is the case, auxiliaries do not govern verbs, allowing them to have a different language index. Joshi (1985) mentioned that auxiliaries cannot be switched because
they constitute closed-class items, and should, therefore, appear in the matrix language (p. 194). However, open-class items, such as participles, can be switched. Azuma (1993) also stated that the language of the participle can differ from that of its preceding auxiliary, as long as the auxiliary’s subcategorization requirements are met (p. 1079).

An issue that the previous studies have in common is that they group all auxiliaries together when referring to their acceptability or unacceptability in codeswitching. It is, therefore, not possible to get at the question of any differences in judgment with respect to codeswitches that include different types of auxiliaries and participles. Nonetheless, a different picture emerges when distinct auxiliaries are considered separately, and this picture becomes clearer in studies that focus on naturalistic production data.

### 1.4 Asymmetry in naturalistic production of codeswitches

As mentioned above, many of the researchers that examine intrasentential codeswitches have collected naturalistic production samples (e.g., Belazi, Rubin, & Toribio, 1994; Di Sciullo, Muysken, & Singh, 1986; Fishman, 1972; Gumperz, 1982; Halmari, 1997; Lipski, 2005; MacSwan, 2000; Mahootian & Santorini, 1996; Myers-Scotton & Jake, 2001; Pfaff, 1979; Poplack, 1980, 1987; Timm, 1975; Zentella, 1997). As a result, these studies provide a rich body of naturally produced oral codeswitches, which reveals the broad distributional codeswitching patterns of different communities and language pairs. Quantitative studies that focus specifically on Spanish-English intrasentential alternations have shown that certain types of syntactic junctures are more prone to undergo switching than others. For instance, switching tends to occur frequently between a subject noun phrase and a verb phrase (e.g., *Los niños* went to school. ‘The kids went to school.’), between a verb phrase and an object noun phrase (e.g., The kids bought *dulces en la tienda*. ‘The kids bought candy in the store.’), and between a main clause and a
subordinate clause (e.g., *Los niños fueron a la tienda* that sells the best toys in town. ‘The kids went to the store that sells the best toys in town.’). Switches between a determiner and a noun (e.g., *Ella se quiere comprar el* textbook for her class. ‘She wants to buy the textbook for her class.’) and switches at prepositional phrases (e.g., *She will return tomorrow por la tarde.* ‘She will return tomorrow in the afternoon.’) have also proved to be common switch points (Lipski, 1978; Pfaff, 1979; Poplack, 1980; Timm, 1975). However, switches involving pronouns and negative elements usually do not occur in naturalistic production data.

### 1.4.1 Asymmetry in naturalistic production of auxiliary phrase switches

Specifically with respect to switches involving the auxiliary phrase, a shared finding among these studies is that, in general, these switches are not very frequent. Timm (1975) mentioned that in her oral and written corpora verb phrases containing auxiliaries and main verbs were usually only produced in unilingual structures (p. 478). In Poplack’s (1980) oral corpus, verb and verb phrase switches represented only 6.9% of the total number of intrasentential switches (n = 851, p. 602). However, when these types of switches do occur in natural bilingual production, there are definite differences between their rates of occurrence, depending on the particular type of auxiliary involved. Although Lipski (1978) stated that codeswitching is usually prohibited in verb phrases containing auxiliaries because they constitute atomic, unbreakable phrasal elements, he added that there are occasional exceptions to this general rule. The example that the author cited as an exception is a switch between the auxiliary *estar* and an English present participle. In fact, in a later study on spontaneous codeswitching samples produced by members of a Mexican-American community in Houston, Texas, Lipski (1985) referred again to auxiliary+participle switches, this time stating that “[o]nly the boundary between *haber*/have and the past participle seems to be a truly insurmountable obstacle to intrasentential shifting” (p. 23).
This different occurrence between estar+English participle and haber+English participle switches was reiterated in Lipski’s general description of Spanish-English codeswitching in the United States (2008, p. 232). Moreover, Pfaff’s (1979) analysis of conversational codeswitching data from 200 Spanish-English bilinguals from California and Texas displayed a predominance of estar+English participle switches over haber+English participle switches. Out of the 45 switches that occurred within the verbal phrase, 16% involved the former type of switch and only 4% involved the latter (p. 299). Three particular constraints that Pfaff proposed can account for the different occurrence of these two types of codeswitches. First, due to the need to express a subject and tense/mood/aspect in verb switching, switches to English verbs that are not morphologically adapted to Spanish can occur only when they are preceded by an inflected Spanish verb or auxiliary. Second, codeswitching is more common at syntactic sites where the linear structure of the two languages coincides. Third, although the word order of a syntactic phrase is structurally parallel in both languages, not all structurally possible switches are equally probable. Therefore, switches between auxiliaries and participles are allowed; however, some switched structures (e.g., estar+English participle switches) may be more likely than others (e.g., haber+English participle switches).

1.4.2 Accounting for the asymmetrical production of auxiliary phrase switches

Although most of the grammatical and psycholinguistic studies on codeswitching cannot explain the difference between estar+English participle switches and haber+English participle switches regarding their appearance in naturalistic production, certain approaches do provide insight on this issue.
1.4.2.1 Morpheme classification account

One proposal that can account for the asymmetrical presence of these two types of codeswitches in production is found in the MLF model and the 4-M model (Myers-Scotton, 1993b; Myers-Scotton & Jake, 2000a, 2000b, 2001). Jake and Myers-Scotton (2009) have shown that just because different lexical items belong to the same lexical category does not mean that they necessarily constitute the same type of morpheme, according to their 4-M model of morpheme classification. Both haber and estar are auxiliaries, under the 4-M model, but they may not constitute the same class of element. Although both are polymorphemic and convey tense and aspect, haber functions as a place holder in Spanish, given that the well-formedness requirements of the language oblige haber to appear in participial structures. In the 4-M model, such place holders are classified as outsider late system morphemes. The lack of production of haber+English participle switches may be due to the fact that haber+Spanish participle is required for the expression of tense and aspect in Spanish. In the switched context, this relationship is violated because of the presence of the English participle, which weakens the expression of the bundled tense and aspect features. By contrast, estar is not a place holder in Spanish. It expresses progressive aspect, it does not have to be strictly adjacent to its participle in the progressive structure, and it can function as a main verb in other verbal structures. Under the 4-M model, estar patterns with early system morphemes, as it has more semantic weight than haber. Therefore, estar is conceptually activated, and switches involving the estar+English participle structure should be evidenced.
1.4.2.2 Grammaticalization account

Another proposal that provides an explanation for the different occurrence in production of these two types of codeswitches comes from the literature on grammaticalization. Grammaticalization is defined as a process of linguistic change whereby autonomous lexical items or items with a less grammatical status gain a more grammatical status (Heine, 2003, p. 575; Lehmann, 2002, p. 10). In some instances, not only a single lexical item, but also an entire construction composed of different lexical items, may come to serve a grammatical function in a particular linguistic context (Hopper & Traugott, 1993, p. 4). Repetition plays a central role in grammaticalization, causing frequently used sequences of words or morphemes to become automated as a single processing unit (Bybee, 2003, p. 603).

The Spanish auxiliary haber comes from the Latin verb habēre, whose original meaning was ‘to have,’ ‘to possess,’ ‘to bring,’ or ‘to hold’ (Coromines & Pascual, 1980). Over time, habēre adopted new uses and it began to appear as an auxiliary verb with accompanying elements. It was used with a preceding infinitive to express obligation and future, and with a past participle to convey anteriority. Habēre also arose as an existential verb and as an impersonal verb. There are several indications of habēre’s grammaticalization process. It became more frequent as it gained more meanings and uses. Since it spread from being used in very specific situations to appearing in broader contexts, it also increased in generality. In addition, habēre lost its semantic transparency as it moved farther away from its original meaning. It went through semantic bleaching, that is, the process of shifting from a very concrete meaning to a more abstract one (Bybee, 2003, p. 605).

During the Middle Ages, with the introduction of tener ‘have’ as a new verb of possession, the use of aver as a main verb decreased, first from prototypical possession to
metaphorical possession², and, then, until it was no longer used to express possession at all (Garachana Camarero, 1997; García Gallarín, 2002). Aver’s uses were even further reduced. In the case of the future auxiliary, aver was grammaticalized until it became an affix to the main verb. In the case of obligation, it was substituted by the modal verb deber ‘should.’ As an impersonal verb, it was replaced by the use of se in impersonal middle constructions. Moreover, in the Middle Ages several simple verbs began to appear more often than periphrastic constructions with aver (e.g., aver alegría ‘to have happiness’ became alegarse ‘to cheer up,’ aver miedo ‘to have fright’ became asustarse ‘to become frightened,’ and aver nombre ‘to have a name’ became llamarse ‘to be named’ (García Gallarín, 2002, p. 20).

In contemporary Spanish, haber has lost all of its original possession meaning as well as most of its other uses. As an autonomous lexical item, it is only employed as an existential verb. In this case, it is used on its own in all tenses except in the present tense, where the originally analytic construction ha+i has become hay ‘there is/are.’ It rarely appears in fixed phrases, such as haber+menester (e.g., he menester ‘(I) need’), ha lugar ‘there is cause for,’ tiempo ha ‘it has been a long time,’ and haber+de+infinitive (e.g., he de cantar ‘(I) should sing’), but these are considered archaic and belong mostly to written language (Alarcos Llorach, 1995, p. 186). Besides these uses, haber only surfaces in combination with the past participle to express anteriority or perfect aspect. Consequently, haber has grammaticalized to a point where it does not exist on its own, but rather coexists with the past participle as a single unit.

² Prototypical possession refers to a situation in which there are very clear differences between the possessor and the possessum: the former is human, agentive, volitional, and has control over the latter, which embodies the opposite traits (it is inanimate, non-agentive, non-volitional, and incapable of having control over anything). With metaphorical possession, some of these categorical distinctions are altered. For example, the possessor can be non-human. Additionally, the possessum can be human, as in the expression of personal relationships, or it can be abstract, in which case the possessum constitutes actions or events, attributes, feelings, states of mind, and opportunities (Garachana Camarero, 1997, p. 217).
The auxiliary *estar* comes from the Latin verb *stare*, which meant ‘to stand,’ ‘to be firm,’ or ‘to remain still and motionless’ (Coromines & Pascual, 1980). This verb’s grammaticalization process was triggered when, in the Middle Ages, it developed from a full verb that expressed the existence of something during a certain time to an auxiliary that, paired with a present participle, described a location and an action which was being performed in an indicated place (Haßler, 2002, p. 170). At this time, nonetheless, each component of the periphrastic form still contributed its own particular meaning to the construction. By the end of the Middle Ages, *estar*+present participle was systematically used to express progressive aspect. In contemporary Spanish, it is also employed in the expression of habitual and frequentative aspects (Torres Cacoullos, 2001, p. 460-461).

Torres Cacoullos (1999, 2001) studied the grammaticalization of periphrastic progressive constructions with three auxiliaries, *estar* ‘be,’ *ir* ‘go,’ and *andar* ‘walk’ or ‘go around,’ by examining their appearance in Old Spanish (12th- to 15th-century) literary texts as well as oral corpora of educated and conversational Mexico City speech. By analyzing the co-occurrence of the three auxiliaries with different classes of main verbs and with different locative and temporal expressions, the researcher demonstrated that these constructions have undergone diverse degrees of grammaticalization, with *andar* representing the newest and least grammaticalized auxiliary, followed by *ir*, and then by *estar*, the most grammaticalized auxiliary (1999, p. 53). Torres Cacoullos also showed that, despite their grammaticalization, these auxiliaries still preserve part of their original lexical meanings: that of general movement meaning in the case of *andar*, that of allative meaning in the case of *ir*, and that of locative meaning in the case of *estar* (1999, p. 27). The author explained that these progressive constructions changed from free lexical combinations in Old Spanish to conventionalized units in contemporary Spanish. At present, there are several indices of the grammaticalization process of these constructions, displayed by means of their reduction of form and function. First, there is a change from variable order of the auxiliary and
participle to a fixed order, in which the auxiliary only precedes the participle. Second, there is a change from positional variation of the object clitic (either before the auxiliary, between the auxiliary and the participle, or after the participle) to the categorical preposing of the clitic to the auxiliary. Third, there is a reduction in the occurrence and the amount of intervening material between the auxiliary and the participle. Finally, there is a decrease in the occurrence of multiple parallel participles with one auxiliary (2001, p. 445).

Among the three auxiliaries that Torres Cacoullos compared, estar proved to be the most grammaticalized. However, if haber is also taken into account, this auxiliary seems to be even more advanced than estar in the grammaticalization process. The use of estar may be more restricted now than it was in Old Spanish, but it is still not as constrained as the use of haber. For example, there are instances, though more literary or poetic, where the participle precedes the auxiliary estar (e.g., cantando estoy ‘singing (I) am’). In addition, although the clitic no longer appears between the auxiliary and the participle, there are occurrences in which the clitic does not necessarily precede the auxiliary, but rather follows the participle (e.g., estoy recogiéndolo ‘(I) am picking it up’). Moreover, there is still occasional appearance of intervening material between estar and its participle (e.g., estoy en casa estudiando ‘(I) am at home studying’) as well as the occurrence of multiple participles with one estar auxiliary (e.g., estoy leyendo y escribiendo ‘(I) am reading and writing’). None of these exceptions, however, are true for the auxiliary haber. In contemporary Spanish, haber is always immediately preposed to the participle and the use of multiple participles with one haber auxiliary is extremely rare and considered archaic (Alarcos Llorach, 1995, p. 149). Furthermore, the object clitic is restricted to precede the auxiliary and there is no intervening material between haber and its participle. Estar can be followed by other expressions, such as adjectival phrases and adverbial phrases (Coromines & Pascual, 1980). However, with the exception of its use as an existential verb, haber does not occur without a past participle (Alarcos Llorach, 1995, p. 186). It seems, therefore, that haber has grammaticalized
further than *estar*. The different degree of grammaticalization of these two auxiliaries can explain their distinct occurrence in naturalistic codeswitching production. *Estar* still retains some of its original lexical meaning and, thus, seems to be more autonomous from its accompanying present participle, allowing a codeswitch to occur between them. By contrast, *haber*, which has lost all of its original meaning and now only exhibits grammatical information, is so linked to its following past participle that both elements have become an indivisible phrasal unit. This degree of fusion or boundedness between *haber* and its participle may explain their resistance to allow for a codeswitch. Instead, they must occur in the same language.

1.4.2.3 Structural distance account

Yet another possible account for the uneven distribution of *estar*+English participle switches and *haber*+English participle switches is understood under a theoretical analysis of the syntactic distance between each of the elements involved in these two constituents. According to the theory of Universal Grammar, *estar*+participle and *haber*+participle constitute distinct underlying syntactic structures. In the case of the progressive structure, the auxiliary *estar* and the present participle belong to separate constituents, as illustrated, in Figure 1-1, for the sentence *Juan está caminando* ‘John is walking.’
Here, the nominal agent *Juan* is generated in the specifier of the verbal phrase and moves up to the specifier of the inflectional phrase in order to check case features. The auxiliary *estar* also moves up to the inflectional phrase in order to check features of tense and agreement. The lexical verb *caminando* remains in the verbal phrase. As a result, in this structure, the auxiliary *estar* and the present participle end up in distinct constituents: the former is in the inflectional phrase and the latter is in the verbal phrase.

The syntactic structure of *haber*+participle is different. Although this interpretation is not undisputed, there are several researchers who consider the Spanish auxiliary *haber* to be a clitic (Suñer, 1987; Zagona, 2002). Some investigators have noted that this is particularly true in the case of monosyllabic auxiliary forms, such as *he* ‘(I) have,’ *ha* ‘(he/she/it) has,’ and *han* ‘(they) have’ (Keniston, 1937; Suñer, 1987). During the course of the derivation, *haber* cliticizes to the past participle, which, therefore, acts as a host. Then, both *haber* and the past participle move up to the inflectional phrase as a unit to check tense and agreement features.\(^3\) The syntactic

\(^3\) This interpretation coincides with ideas presented in the Minimalist Program. Within this framework, all compound verbal structures are individual lexical entries and, therefore, not subject to the insertion of elements between the two parts of the lexical entry.
configuration of perfect structures is displayed in Figure 1-2 for the sentence *Juan ha caminado* ‘John has walked.’

Regardless of the reason for the distinct appearance of these two types of switches in production, for purposes of this dissertation, the crucial finding is the documented production asymmetry with respect to alternations involving the auxiliary phrase. Specifically, switches between the Spanish auxiliary *estar* and the English present participle are comparatively more frequent than switches between the Spanish auxiliary *haber* and the English past participle. Table 1-1, which includes examples drawn from several Spanish-English codeswitching corpora, further displays the uneven distribution between the two types of switches.\(^4\)

\(^4\) In all examples included in this dissertation, the elements of codeswitched verbal phrases are underlined.
1.5 Comprehension of codeswitches

The preponderance of switches involving the progressive structure over those involving the perfect structure demonstrates their different behavior in bilingual production. Although research has examined the production of these codeswitches, the way in which Spanish-English bilinguals comprehend them has been largely unexplored. Clearly, codeswitching is not only a production phenomenon. All codeswitched structures that are produced by a bilingual must, in turn, be processed by the interlocutor’s comprehension system. Unlike codeswitched production, which is under the control of the bilingual speaker, during comprehension, codeswitches can be unexpected for the bilingual interlocutor and, potentially, more difficult to process than
unswitched language. With this idea in mind, some researchers have recently begun to examine the comprehension of written codeswitches, particularly by means of electrophysiological responses (e.g., Moreno, Federmeier, & Kutas, 2002; Proverbio, Leoni, & Zani, 2004). In fact, in Proverbio et al.’s (2004) study, Italian-English interpreters displayed increased processing costs with codeswitched sentences, compared to unilingual sentences. That is, when the sentence-final words were switched, they elicited larger N400 responses than did the sentence-final words that were presented in the same language as the rest of the sentences (p. 1649). In Moreno et al.’s (2002) study, English-Spanish bilinguals also exhibited greater processing costs with codeswitched sentences than they did with unilingual sentences. In this case, the processing effect of codeswitches was not displayed in N400 responses, which are considered to reflect semantic integration, but instead in Late Positive Complex (LPC) responses, which are associated with processing of an unexpected or improbable event. Specifically, in this study, participants’ LPC response was larger when they encountered a sentence-final noun that was switched than when it was unswitched (p. 202).

Importantly, in these studies the codeswitching pattern to which participants have been exposed has consisted of the insertion of a final word (generally a noun) in one language in an otherwise unilingual sentence in the other language (e.g., Each night the campers built a fuego [‘fire’], Moreno et al., 2002, p. 191). These single noun insertions, although very common in Spanish-English codeswitching production data, are not considered to entail the same degree of linguistic interaction nor to require the same level of proficiency in both languages as the auxiliary phrase intrasentential switches discussed above (Poplack, 1980, p. 605). It is possible that intrasentential switches are processed differently from single-word switches, particularly in the case of more balanced bilingual codeswitchers. In fact, Moreno et al. (2002) provide evidence that suggests this. In their study, the size of the LPC response was modulated by proficiency in the bilinguals’ second language, Spanish. Those who were more proficient in
Spanish (who also tended to report more frequent codeswitching practices) displayed a smaller amplitude LPC response, therefore demonstrating that the codeswitches were less unexpected and less difficult to process for them (p. 202). Given this gap in the literature, the set of studies in this dissertation examines the processing costs involved in the comprehension of intrasentential switches and, therefore, tests the relation between production patterns and comprehension costs of codeswitching by Spanish-English bilingual codeswitchers who are proficient in both languages.

1.6 Relating comprehension costs and production patterns

Psycholinguistic studies examining monolingual linguistic behavior have shown that production and comprehension involve many of the same representations and processes. During comprehension, listeners map the signal onto lexical entries whose morphosyntactic and semantic information becomes available to give structure and meaning to the utterances. Similarly, in production, speakers select lexical items, each carrying syntactic, semantic, and morphological features, which affects the selection of additional words (Treiman, Clifton, Meyer, & Wurn, 2003). Given this connection between both systems, studies have examined how comprehension and production interact and affect each other. Many have shown that there is a very strong correspondence between the distributional patterns of different linguistic structures as found in production and the costs incurred during comprehension of those structures (e.g., Garnsey, Pearlmutter, Myers, & Lotocky, 1997; MacDonald, Pearlmutter, & Seidenberg, 1994; MacDonald & Seidenberg, 2006; Tanenhaus & Trueswell, 1995; Trueswell & Kim, 1998; Trueswell, Tanenhaus & Kello, 1993; Wilson & Garnsey, 2009).

For instance, Trueswell and Kim (1998) and Wilson and Garnsey (2009) conducted studies aimed at discovering how recognition of a verb and the parallel activation of its subcategorization preferences could impact the commitments readers make when resolving
temporary syntactic ambiguity. Using the self-paced reading technique as well as the eye-tracking method, participants read sentences on a computer screen and answered comprehension questions. The sentences contained ambiguous verb complements, that is, verb complements that could be the direct object (DO) of the main verb or the subject of a sentential complement (SC). Both studies found that the processing of verbal structures is influenced by verb bias. In other words, the verbal structure was harder to process when the verb was followed by a complement that did not match the verb’s DO or SC subcategorization preferences, as they have been reported in production.

Further evidence for the strong connection between production and comprehension was supplied in a recent study by Gennari and MacDonald (2009). In this case, the researchers examined the production patterns of relative clauses with respect to their head nouns and the verbs that they include. The results displayed that, in both natural and controlled production, the tendency was to produce passive relative clauses with theme-experiencer verbs and animate head nouns, while active object relatives were preferably produced with agent-theme verbs and inanimate heads. Then, two comprehension experiments found that nouns and verbs occurring in relative clause structures that speakers do not usually produce are, in fact, difficult to comprehend. That is, the verb and head noun types that speakers tend to produce in active or passive relative clause structures are easier to process when they appear precisely in those syntactic structures.

Taken together, these findings demonstrate the importance of construction frequency when it comes to sentence comprehension. They suggest that usage and exposure-based factors guide readers’ initial processing. In turn, they lend support for the Production-Distribution-Comprehension (PDC) model (Gennari & MacDonald, 2009; MacDonald & Thornton, 2009). As clearly displayed by its name, this account links the production system to the comprehension system by means of the language’s distributional patterns. According to this model, the
properties and mechanisms of the production system influence the structure choices that individuals make during production. These decisions create robust distributional patterns for different linguistic structures in the language, which are learned over time by those who are exposed to this input. Then, the distributional patterns become the probabilistic constraints that guide comprehension in a constraint-based system.

In this way, the PDC account is consistent with other constraint satisfaction accounts of sentence processing (MacDonald & Seidenberg, 2006), in which sentence processing is not modular nor is it controlled solely by structural information. Instead, during comprehension, information provided by diverse linguistic sources, such as phonology, morphology, semantics, pragmatics, and syntax, is simultaneously taken into consideration to discard some possible analyses and support others. Thus, multiple (bottom-up and top-down) constraints interact at the same time to determine initial syntactic commitments. The PDC account and other constraint satisfaction models grant a central role to frequency. Within these models, structural regularities in individuals’ experience are an invaluable source of information that influences comprehension processes. In other words, frequent linguistic structures are more readily activated than less common structures, and are, therefore, easier to process during comprehension.

The strong dependence between the production and comprehension systems has only been examined with monolingual participants and unilingual sentences. However, the bilinguals and multilinguals in the world greatly outnumber the monolinguals. In order to adequately characterize the way in which individuals comprehend sentences, it is necessary to examine not only how monolinguals process language, but also how bilinguals (and multilinguals) negotiate the presence of two (or more) languages in a single mind. This dissertation attempts to extend the examination of the correspondence between production and comprehension to the bilingual context, not only by observing how bilinguals process language in general, but by observing how they comprehend codeswitched sentences.
1.7 Justifying a focus on reading comprehension

The set of studies in this dissertation focuses particularly on the reading comprehension of codeswitches. There are numerous differences between speech and writing, including the universality of speech over writing, differences in acquisition, and issues pertaining to message retrievability, prestige, and standardization (Daniels & Bright, 1996). Unlike speech, the teaching of written language generally promotes the belief that some variants of written language represent the “correct” language. In this respect, there has been a tendency to regard codeswitching as almost exclusively an oral phenomenon and to discount written codeswitching as artificial or qualitatively different from oral discourse. However, research shows that the types of codeswitches that occur in written discourse closely match those reported for spoken codeswitching. Importantly, Spanish-English codeswitching is common in written language and, thus, in bilinguals’ reading experience as well. This is particularly true of email interactions, as attested in the extensive written codeswitching corpus of Montes-Alcalá (2005a, 2005b, 2005c). In addition, codeswitching also occurs recurrently in literary works. Callahan (2002) analyzed a corpus of thirty texts (2,954 pages) of short stories and novels containing Spanish-English and English-Spanish codeswitches and found very little variance between the types of codeswitches used in the written texts and in those used in oral production. Table 1-2, taken from Dussias, GuZZardo Tamargo, Valdés Kroff, and Gerfen (in press) compares Montes-Alcalá’s (2005a, 2005b, 2005c) Spanish-English codeswitching data, gathered from email messages between bilinguals, with published oral data by Klavans (1985), Lance (1975), Lipski (1985), Pfaff (1979), and Poplack (1980).
As seen in Table 1-2, despite differences between spoken and written language, the written codeswitching corpora exhibit codeswitch types that are strikingly analogous to those found in natural speech. Therefore, by extension, it is reasonable to expect that the reading comprehension costs of codeswitched sentences should resemble those that would result from auditory processing of the same codeswitched utterances.

<table>
<thead>
<tr>
<th>Codeswitch types</th>
<th>Email interactions</th>
<th>Oral production data</th>
</tr>
</thead>
<tbody>
<tr>
<td>at clausal boundaries</td>
<td>I am sure he’ll appreciate it, aunque le duela ‘I am sure he’ll appreciate it, even if it hurts him’</td>
<td>Aquí con Hope, I am trying to figure out what’s wrong… ‘Here with Hope, I am trying to figure out what’s wrong…’</td>
</tr>
<tr>
<td>at the verb-object boundary</td>
<td>Laura está escondiendo everyone’s drinks! ‘Laura is hiding everyone’s drinks!’</td>
<td>Los hombres comieron the sandwiches ‘The men ate the sandwiches’</td>
</tr>
<tr>
<td>following a preposition</td>
<td>Yo estoy de acuerdo con the group nomination ‘I agree with the group nomination’</td>
<td>Ben viene con the sports news ‘Ben comes with the sports news’</td>
</tr>
<tr>
<td>at the adjective</td>
<td>Pues aquí estoy muy bored ‘Well, here (I) am very bored’</td>
<td>Mi papá es muy protective ‘My dad is very protective’</td>
</tr>
<tr>
<td>at adverbial phrases</td>
<td>Y además, me llamaron para una entrevista at that place ‘And besides, (they) called me for an interview at that place’</td>
<td>Uno no podía comer carne everyday ‘One could not eat meat everyday’</td>
</tr>
<tr>
<td>noun phrase insertions</td>
<td>Espero que no trabajes demasiado, es un three day weekend… ‘(I) hope that (you) do not work too much, it’s a three day weekend…’</td>
<td>Me lo dijo el same night ‘(S/He) told me the same night’</td>
</tr>
<tr>
<td>following a complementizer</td>
<td>Después me quedé pensando por lo que dijiste que that would make the other guy mad ‘Afterwards (I) kept thinking about what (you) said that that would make the other guy mad’</td>
<td>Se me hace que I have to respect her porque… ‘(I) think that I have to respect her because…’</td>
</tr>
</tbody>
</table>
1.8 Relating reading and auditory sentence comprehension

As with unilingual communication, most codeswitching occurs in the spoken domain. Nevertheless, there is a broad consensus regarding the strong connection between auditory processing and reading comprehension. In both reading and listening, sentence comprehension unfolds rapidly over the course of perception, and structure-building is influenced by similar phonological, syntactic, semantic, and referential factors (Tanenhaus & Trueswell, 2007). Studies have shown that reading activates much of the same system employed in auditory language processing (Perfetti, 1994). For instance, Fodor (1998) proposed that syntactic processing during reading proceeds through subvocal phonological encoding of prosody, which also guides auditory comprehension. Steinhauer and Friederici (2001) confirmed these findings by showing that language processing in both reading and auditory comprehension is reflected in the same event-related potential (ERP) component known as the Closure Positive Shift. Moreover, in another ERP study, Osterhout and Holcomb (1993) found that hearing words in a sentence that are inconsistent with the preferred syntactic structure produced a P600 effect similar to that elicited during reading (Osterhout & Holcomb, 1992). This suggests that processing decisions during sentence comprehension are constant across both modalities, at least in some circumstances.

Work using the visual world paradigm (Allopenna, Magnuson, & Tanenhaus, 1998) also provides converging results for written and auditory processing. Using this eye-tracking technique, Snedeker and Trueswell (2004) found that despite the presence of potent visual cues to the visual referential context, participants remain sensitive to verb bias information during auditory processing, a result consistent with the ambiguity resolution findings in reading studies (e.g., Britt, 1994; Garnsey, Pearlmutter, Myers, & Lottoky, 1997; Trueswell, Tanenhaus, & Kello, 1993). Similarly, a self-paced reading study by Trueswell and Kim (1998) and a spoken-
language comprehension experiment by Novick, Kim, and Trueswell (2003) tested how the fast priming of verbs affects processing of ambiguous sentences for readers and listeners, respectively. In both studies, the syntactic preferences of briefly presented prime words modulated parsing decisions. In the reading comprehension study, participants read sentences that included direct object/sentential complement ambiguity, although they all ended with sentential complements (e.g., *The photographer accepted the fire could not be put out*). If the main verb of the sentence (i.e., *accepted*) was primed with a verb that tends to be used with a sentential complement (e.g., *realized*), participants exhibited less processing costs in the disambiguating region (i.e., *could*). If, instead, the main verb was primed with a verb that tends to be used with a direct object (e.g., *obtained*), participants displayed increased processing difficulty in the disambiguating region, where they encountered the sentential complement of the main verb (Trueswell & Kim, 1998, p. 115). In the auditory comprehension study, participants heard sentences (e.g., *Turn the doll with.*) while viewing a scene with objects. The sentences ended with either a noun phrase modifier (e.g., *the pink dress*) or a verb phrase instrument (e.g., *the stick*). Participants looked at the instrument in the referential scene (i.e., a stick) more often when they heard sentences that included an instrument-bias verb prime (e.g., *clean*) than when they heard sentences that included a modifier-bias verb prime (e.g., *hug*; Novick, Kim, & Trueswell, 2003, p 71). Taken together, these results show that similar lexically specific parsing processes take place in auditory language comprehension and in written language comprehension.

1.9 Task effects on sentence processing

Psycholinguistic studies that examine sentence comprehension make use of different secondary tasks. In these studies, participants complete a main task that consists of listening to a sentence or reading a sentence. After processing the sentence, participants are usually asked to
complete a secondary task to ensure that they focus on the main task and pay attention to the sentences that they hear or read. Some researchers have found that the particular secondary task that participants are asked to complete heavily influences the way they process sentences. For instance, Williams (2006) examined the processing difficulties that native and nonnative speakers of English incurred as they read wh-questions with long distance dependencies. The wh-questions constituted adjunct extractions and included a potential gap after the verb, whose filler was either plausible as the direct object (e.g., *Which car did the man buy the really expensive radio for two months ago?*) or implausible (e.g., *Which friend did the man buy the really expensive radio for two months ago?*; p. 87). The second goal of the study was to examine potential task effects on participants’ reading times. Participants read sentences on a computer screen one word at a time and their reading times on each word were recorded. In addition to reading, they completed one of two secondary tasks. The native and nonnative speakers in Experiment 1 performed an on-line plausibility judgment on each sentence by means of the stop-making sense task, in which they were asked to press a computer key as soon as they thought that a sentence stopped making sense. Those in Experiment 2 read comprehension/memory probes, which included a missing word, after every pair of sentences and completed them orally.

Results showed that the nature of the task affected participants’ reading times. In Experiment 1, both native and nonnative speakers processed sentences similarly by making use of plausibility constraints in the same way. In other words, both groups of participants displayed longer reading times in the immediate post-verbal region in the Plausible-at-Verb condition compared to the Implausible-at-Verb condition, reflecting the increased difficulty of revising the initial structural hypothesis when that interpretation was plausible. In Experiment 2, both native and nonnative speakers behaved differently from each other and from their respective counterparts in Experiment 1. Only those native speakers with better performance on the comprehension/memory task displayed a plausibility effect in the immediate post-verbal region in
the Plausible-at-Verb condition; however, this effect was larger towards the end of the sentence.

The native speakers with lower scores on the comprehension/memory task showed some evidence of a plausibility effect at the post-verbal noun, but the effect disappeared towards the end of the sentence. Amongst the nonnative speakers, only those with high scores on the comprehension/memory test showed any plausibility effects, but these were delayed until the final words of the sentence (p. 81).

Williams (2006) concluded that incrementality of interpretation depends on general cognitive factors, such as working memory and motivation. When the task required clear attention to meaning (i.e., the stop-making sense task), both groups of participants evaluated plausibility incrementally and showed early plausibility effects. However, without this requirement (i.e., in the comprehension/memory task, which better resembles a normal reading situation), the effects were delayed, and even eliminated in the case of some participants because the reading situation did not encourage incremental interpretation, but instead simply motivated levels of comprehension that were sufficient for completing comprehension/memory probes. In sum, semantic processing appears to be task-dependent and participants display the capacity to alter their reading strategy within a particular task (p. 86).

Task effects also emerged in Macizo and Bajo’s (2006) study with Spanish-English professional translators and untrained bilinguals. In this case, the researchers tested two assumptions of the vertical and horizontal theories of translation: one regarding the involvement of working memory in reading and translating and another regarding the serial/parallel nature of comprehension and reformulation. Both groups of participants read Spanish sentences for repetition or for translation. In order to test these assumptions, the sentences were manipulated for lexical ambiguity and for memory load. Half of the sentences included a lexically ambiguous target word and the other half included a lexically unambiguous target word. High memory load
sentences incorporated more words between the target word and the disambiguating word than low memory load sentences.

The vertical view assumes that comprehension of the source language and access to the target language are independent processes that are carried out in a serial manner. If this is true, reading for translation and reading for repetition should impose similar demands on working memory and, thus, display similar reading times. In addition, ambiguity and memory load should have equivalent effects when reading for repetition and reading for translation. According to the horizontal view, reading for translation is more demanding than reading for repetition because reading the source text engages partial reformulation processes. If this is the case, then reading for translation should be more costly than reading for repetition. Moreover, ambiguity and memory load should play an important role in reading for translation, but their role should be diminished when reading for repetition.

Results showed that, overall, both groups of participants displayed the same pattern of lexical effects (i.e., an interaction of lexical ambiguity and memory load) on their reading times. The instructions that participants received clearly influenced the processes involved in reading and understanding visually presented sentences. During the repetition task, participants were not affected by the presence of lexical ambiguity in the sentences. In contrast, during the translation task, they showed lexical ambiguity effects with high memory load sentences. When participants were asked to read the sentences for translation, they engaged processes that exhausted more working memory resources than those that are needed for reading for repetition, a finding that lends support for the horizontal theory of translation (p. 14).

Task effects have also been found in studies using ERP measures of sentence processing (Hahne & Friederici, 2002; Huddleston et al., 2003; Wang, Ditman, Choi, & Kuperberg, 2010). Hahne and Friederici (2002) used ERPs to investigate the time-course of semantic and syntactic processes in auditory language comprehension. They were also interested in evaluating the
impact of task demands on sentence comprehension strategies. The critical items included correct sentences, sentences that contained a semantic violation (semantic condition), sentences that contained a phrase structure violation (syntactic condition), and sentences that contained both types of violations (combined condition). In Experiment 1, participants were asked to judge the overall correctness of the sentences that they heard. That is, the instructions did not make distinctions between the different types of errors that participants encountered. Experiment 2 required participants to judge the sentences only for their semantic coherence, namely to decide whether the sentence made sense or not, and to disregard phrase structure violations.

Once again, results demonstrated an effect of task manipulation, but only with certain ERP components. The instructions clearly influenced the later components, but not the earlier anterior negativity. In Experiment 1, the semantic condition elicited an N400 effect and the syntactic and combined conditions elicited an early anterior negativity followed by a P600 effect. In Experiment 2, the combined condition elicited an early anterior negativity and an N400 effect that resembled the magnitude of the N400 effect elicited in the semantic condition. Furthermore, the syntactic condition elicited early anterior negativity, as well as N400 and P600 effects similar to those observed for the correct sentences (Hahne & Friederici, 2002, p. 352). The researchers concluded that late ERP components, such as N400 and P600, reflect controlled processes and are dependent on task demands whereas initial parsing processes associated with earlier negativity are independent of semantic constraints and task requirements.

The studies discussed in this section have shown that different tasks or instructions engage distinct processing strategies. Many of the codeswitching studies discussed in the previous sections made use of acceptability judgment tasks. It is possible that the codeswitching patterns that emerged from these studies were simply an artifact of the metalinguistic task itself. The judgments of acceptable and unacceptable switches provided by participants may not reflect—and, as seen from the findings described in sections 1.3 and 1.4, in fact, do not necessarily
reflect– the types of codeswitches that bilinguals engage in natural speech. Data obtained by means of acceptability judgment tasks may, therefore, need to be re-evaluated. As mentioned above, the main task in the present set of studies is a reading task. After completing the main task, participants are asked to complete two secondary tasks: an acceptability judgment task and a comprehension task. This manipulation allows an examination of the influence of task on participants’ reading patterns.

1.10 Aims of this dissertation

This dissertation has two main objectives. First, it aims to test whether there is a correspondence between production and comprehension costs of codeswitched expression. That is, are frequently produced codeswitches easier to process than less commonly produced codeswitches? In this case, two types of Spanish-English auxiliary phrase codeswitches are compared. Despite their superficial similarities, one type of switches (i.e., estar+English participle switches) is more predominant in natural codeswitching corpora than another type of switches (i.e., haber+English participle switches), leading to the assumption that Spanish-English bilinguals are presumably more often exposed to the former type of switches, compared to the latter. The goal, thus, is to see whether frequency of exposure to these codeswitched structures can modulate bilinguals’ comprehension difficulty. This would extend findings from the literature on sentence processing, which have only been examined in a unilingual context, to the codeswitching context. If this is the case, then processing of codeswitched language could also be said to conform to constraint satisfaction models of language processing, such as the PDC model (Gennari & MacDonald, 2009), discussed in section 1.6. Importantly, factors beyond the structural properties of language, such as frequency of exposure to particular codeswitched
structures, would be considered to play an important role in the comprehension of those structures.

The second objective of this dissertation is to examine how a particular task influences the bilinguals’ comprehension patterns. Two tasks are compared: an acceptability judgment task, which has been extensively used in theoretical codeswitching studies, and a comprehension task, broadly used in psycholinguistic studies on language processing. This will provide evidence as to whether both of these tasks generate comprehension costs that reflect the tendencies found in naturally-occurring codeswitched production. Finally, the studies in this dissertation also examine the role of age of second language acquisition in the way that Spanish-English bilinguals process these two types of codeswitches.

The remainder of this dissertation is organized in the following manner. Chapter 2 describes a corpus study of estar+English participle switches and haber+English participle switches in both oral and written corpora. Chapter 3 presents the results of an eye-tracking study, conducted in State College, Pennsylvania, that examines the comprehension costs associated with these two types of auxiliary phrase switches. Chapter 4 discusses a replication of the eye-tracking study with bilinguals who belong to an established community of codeswitchers in West Harlem, New York. Chapter 5 comprises a study of the switch costs involved in the comprehension of codeswitches. Chapter 6 concludes with a summary of the findings, implications, and future directions.
Chapter 2

Spanish-English auxiliary phrase switches in oral and written corpora

Early studies on the naturalistic production of Spanish-English codeswitches provide information regarding the general codeswitching production patterns of Spanish-English bilinguals. Based on the data reported in these studies, it is evident that certain types of codeswitches are produced more frequently than others. Specifically with respect to codeswitches involving the auxiliary phrase, studies have shown that switches between the Spanish auxiliary *estar* ‘be’ and the English present participle occur more often than switches between the Spanish auxiliary *haber* ‘have’ and the English past participle (Callahan, 2004; Lipski, 1978, 1985; Pfaff, 1979; Poplack, 1980). However, the reported predominance of switches involving the progressive structure is limited in that it is based on the documentation of very few examples, mainly those included in Table 1-1 of Chapter 1. Since auxiliary phrase codeswitches are not one of the most frequent types of Spanish-English codeswitches overall, they have not been highlighted in previous works, even less so in publications on the written production of codeswitches. Therefore, additional inspection of examples of switches involving the progressive and perfect structures is necessary in order to verify the production patterns described in these earlier studies. This will strengthen the claim that *estar*+English participle switches are indeed more frequent than *haber*+English participle switches in contemporary Spanish-English codeswitching production. With this purpose in mind, an examination of oral and written Spanish-English codeswitching corpora was performed, directed specifically towards the search for switches involving the auxiliary phrase.
2.1 Data

2.1.1 Oral corpus

The oral data used in this study consist of recordings of informal conversations between pairs of Spanish-English bilingual speakers living in Miami, Florida. The corpus was compiled between February and April 2008 by members of the ESRC Centre for Research on Bilingualism in Theory and Practice. The researchers collected the data by providing participants with a portable digital recorder and asking them to record an informal conversation between themselves and another speaker of their choice, with whom they were familiar. In most cases, the conversation partner for each participant was a family member, friend, or co-worker. In order to minimize the effects of the Observer’s Paradox, the researchers left the room where the recording took place. The recorded conversations lasted approximately 30 minutes. Personal information was gathered from participants by means of a background questionnaire. Bilingual proficiency was measured with participant self-assessment rates. The researchers from the ESRC Centre transcribed the recordings with the CHAT transcription system and its associated software CLAN (MacWhinney, 2000). The transcriptions were checked for quality and reliability by other members of the ESRC Centre, also trained in CHAT and CLAN. The average reliability score for the Spanish-English Miami corpus was 94%. After final checks, the transcriptions were made available on Talkbank (http://talkbank.org/browser/index.php?url=BilingBank/Bangor/Miami/).

A detailed description of the corpus and the methods of recruitment, data collection, and data transcription can be found in the article by Deuchar, Davies, Herring, Parafita Couto, and Carter (2012). Although I did not have direct contact with the bilinguals who participated in this project, I am quite familiar with the oral corpus. I had the opportunity to collaborate with the Corpus-Based Research Group, headed by Dr. Margaret Deuchar, to help transcribe and proofread this
Spanish-English corpus. I also participated in bi-monthly video-conference meetings with the ESRC research group to discuss the CHAT transcription conventions, as well as any issues that arose during the course of the transcription process. Because of this experience, I am aware of the transcription procedures and the coding rules.

For this corpus study, I examined the 26 transcriptions of the Spanish-English Miami corpus that were available on Talkbank as of December 2011. The size of these 26 transcriptions is approximately 390,000 words. Two speakers take part in each of the transcriptions. Overall, the speakers in the transcriptions vary with respect to sex, age, occupation, nationality, and level of proficiency in each of their languages. Twenty male speakers and 32 female speakers participate in the conversations. Their mean age is 32 (range 9-66). Some of the most common occupations among the speakers are: student, teacher, manager, secretary, and clerical worker. Moreover, their reported nationalities are: American, Cuban, Colombian, Dominican, Ecuadorian, Jewish, Nicaraguan, Puerto Rican, Venezuelan, or a combination of any of these. The topics of conversation include the speakers’ daily activities and happenings, as well as discussion on current social and political events.

### 2.1.2 Written corpus

The written data comprise an editorial column entitled La Calentita: Gibraltar’s National Dish, which is included in the online version of the Gibraltar newspaper, Panorama (http://www.panorama.gi/). The Calentita column has a weekly appearance in the newspaper. The entries simulate a friendly conversation between two fictitious characters, Cynthia and Cloti, who comment on local socioeconomic and political events in a humorous tone. In accordance with their conversational nature, the entries are written in informal language and include numerous instances of Spanish-English codeswitching. The written corpus is produced by one
columnist; therefore, it may seem far-fetched to use these data to generalize to the codeswitching production of other bilingual Gibraltarians. However, the codeswitching patterns found in the *Calentita* column have proven to coincide with those of fluent bilingual Gibraltarians (Moyer, 1995). Consequently, the codeswitching patterns presented in the column can be considered to be consistent with those of the codeswitching community in Gibraltar. For this corpus study, I examined 88 entries of *La Calentita: Gibraltar’s National Dish*, published between March 2004 and November 2011. The size of the written corpus is approximately 25,300 words.

2.1.3 Differences between the oral and written corpora

There exist clear differences between speech and writing because they constitute different modes of expression. Some of the basic differences have to do with the age, use, and acquisition of these two modes of expression. Speech is older than writing. Moreover, speech is universal, but writing is not. In addition, while people start to speak at a very early age in life, learning to write typically builds on learning to speak (Daniels & Bright, 1996, p. 1). Another difference between these two modes of expression relates to their levels of structure. Speech consists of phonemes or units of sound that combine to form meaningful morphemes. There is more variability with respect to writing. For instance, alphabetic scripts consist of graphemes that form morphemes; syllabary scripts are based on spoken syllables as the basic unit; with logographic scripts, a character corresponds to a morpheme (Daniels & Bright, 1996, p. 4). Differences between speech and writing also include issues of prestige, standardization, formality, and change. Although there are different dialects of spoken languages, which vary across geographical areas and social groups, the written norm is standardized by governmental, educational, and literary institutions. Because written language is associated with these institutions, it is generally considered prestigious and it is associated with more formal style.
Finally, due to its permanence and standardization, written language changes at a slower rate than speech (Daniels & Bright, 1996, p. 2).

In addition to the general differences between speech and writing, the two corpora analyzed here differ in several fundamental ways. First, they belong to distinct speech communities. Therefore, the corpora present differences regarding the Spanish and English varieties used by the participants. The speakers in the oral recordings use United States (U.S.) English mixed with different varieties of Caribbean, Central American, or South American Spanish. The *Calentita* column writer, on the other hand, portrays speakers of British English and Peninsular Spanish. The particular Spanish variety spoken in Gibraltar is strongly influenced by southern Peninsular Spanish, specifically from Algeciras and Cániz. Referred to as Yanito, this variety is described as an Andalusian Spanish-dominant form of expression that integrates lexical and syntactic elements from English, and, to a lesser degree, from other languages, such as the Genoese Italian dialect (Levey, 2008, p. 3; Lipski, 1986, p. 416-417).

The speech communities also differ with respect to the public and private distribution and use of each language. Gibraltar, located at the tip of the Iberian Peninsula, has been a self-governing British overseas territory since the early 1700s. Despite its official condition as a British territory, many of Gibraltar’s residents are of Spanish descent. Although they maintain strong cultural ties with Spain, Gibraltarians strongly oppose Spain’s occasional claims of the territory along with any proposal of shared sovereignty (Lipski, 1986, p. 415). The current political situation of Gibraltar has left its residents with ambivalent feelings towards both Spain and England. They often perceive themselves as second-class British citizens and report feeling looked down on by mainland British citizens and ignored or forgotten by Spain (Lipski, 1986, p. 415; Moyer, 1995, p. 222). In everyday informal activities with family and friends, Spanish is more commonly used than English, even in the case of some British residents without Spanish descent. Although Spanish is the most widely spread language in Gibraltar, English is
predominant in specific areas of Gibraltarian life, such as government, education, religion, and relations with the military (Lipski, 1986, p. 416; Moyer, 1995, p. 223). It also constitutes the language of the media. With the exception of the Calentita editorial notes in Panorama, Spanish is not used in any Gibraltarian newspapers. Moreover, all television and radio shows are solely in English, and English is the language used for general instruction in schools. Despite the fact that most of the students speak Spanish as a native language, it is only offered in schools as an optional foreign language course (Lipski, 1986, p. 416).

The situation in Miami is somewhat different. In this case, Hispanics and Latinos make up approximately 70% of the entire Miamian population (U.S. Census Bureau, 2010a). Although they reside in different sections and neighborhoods of the city, the Latino population, mostly from Central America and the Caribbean, tends to concentrate in particular Hispanic enclaves on the northern and western sides of Miami. The number of Spanish-speaking individuals is so large in Miami that it has been recognized as the primary language in the city and it is used in both private and public domains. Although English is the language used in government, Spanish is widely present in the media. There are several major Spanish-language newspapers, including El Nuevo Herald and Diario Las Américas. Miami also offers numerous television and radio shows in Spanish. It is the headquarters and main production city of many of the world’s largest Spanish television networks, such as Telemundo, TeleFutura, Galavisión, and Univisión. Although, English is the official language used in most schools, the Miamian students represent the largest minority public school system in the U.S., with 60% of the students being of Hispanic origin (SchoolWorks & The Broad Prize for Urban Education, 2007). Consequently, there are several schools in the city that offer bilingual education.

Despite the differences between both speech communities regarding the language varieties and the domains of use of each language, both groups share similarities with respect to their Spanish-English bilingual behavior. In both cases, bilingualism brings about a wide range
of phenomena, including lexical borrowings and semantic and syntactic calques. More importantly, the bilingual speakers of Gibraltar and Miami seem to display analogous patterns of codeswitching (compare Moyer, 1995, for Gibraltarian codeswitching patterns with Lipski, 1985; Pfaff, 1979; Poplack, 1980, for U.S. codeswitching patterns).

2.2 Extraction and coding methods

For the analysis, I searched the transcriptions and the editorial column entries for instances of the progressive structure (i.e., the progressive auxiliary in its full or contracted form followed by a present participle) and the perfect structure (i.e., the perfect auxiliary in its full or contracted form followed by a past participle). The entire sentence that included each instance of the progressive or perfect structure was selected. In the written corpus, sentences were divided by periods. In the oral corpus, there were instances in which the transcriptions lacked punctuation that clearly distinguished sentences. In these cases, the entire speaking turn of a participant was considered a sentence. If the sentence or conversational turn included more than one instance of the progressive or perfect structure, each instance was considered separately. Therefore, sometimes the same sentence was extracted more than once. All grammatical subjects for the progressive and perfect structures were included in the analysis (i.e., noun phrases and pronouns, all grammatical persons, animate and inanimate subjects, individual and collective subjects), as well as all types of sentences (i.e., declarative and interrogative, affirmative and negative). Moreover, the progressive and perfect structures were included with auxiliary forms in the present, past, and future tenses and the indicative and subjunctive moods.

After the initial extraction from both corpora, instances of the progressive and perfect structures were excluded for several reasons. Tokens of the progressive structure were excluded if the structure was used to refer to a future action. The use of the progressive structure to signal
a future action only exists in English; in Spanish the future action is expressed with the simple future or the periphrastic future tenses. This particular use of the progressive structure could be considered an environment in which a codeswitch might be blocked; however, codeswitches of this type are not completely unheard-of in Spanish-English codeswitched production (Lipski, 2008, p. 232). In any case, because this dissertation examines progressive structures used to express progressive aspectual meaning, all instances of the progressive structure used to express a future action were excluded from further analysis. In the case of the oral corpus, tokens were also eliminated from the analysis if the auxiliary was not pronounced and, therefore, not spelled out in the transcriptions. Moreover, tokens were excluded when they contained the perfect structure with the past participle of get because these phrases had a meaning of possession instead of a past action meaning. Finally, tokens that comprised idioms or fixed phrases with either the progressive or the perfect structure were eliminated from further analysis. The following four examples illustrate each type of exclusion mentioned above.

(9) Progressive structure with future meaning
    I’m leaving on Friday. (CHL, herring1)

(10) Unpronounced auxiliary
    because they offering him a big salary (JUL, zeledon5)

(11) Perfect structure with past participle of get
    and then he’s got this black and blue bruise (ANT, zeledon11)

(12) Idiom
    You are telling me, Cynthia de mi corazón [‘of my heart’], but it would not have been that bad had he not come with the Andalusian flag, which he displayed in Convent Place for all to see.5 (Calentita, 15/2/2008)

A total amount of 36 tokens were excluded from the analysis, representing less than 3% of the data. From the resulting tokens, I selected only those that contained a codeswitch somewhere in the sentence, in other words, sentences that included both Spanish and English

5 You’re telling me! is an idiom that is used in a situation where something is so clear that it does not need to be said (http://www.idiomconnection.com/xyzquiz.html).
words. Therefore, I excluded all unilingual English or unilingual Spanish sentences. These codeswitched sentences constituted the codeswitched corpus for the analysis. For each sentence or conversational turn, I coded where a codeswitch occurred with respect to the progressive or perfect structure by noting whether the codeswitch occurred immediately before or immediately after the auxiliary phrase, at the auxiliary, at the participle, or elsewhere in the sentence. There were occasions in which multiple codeswitches occurred in a single sentence or conversational turn. In these cases, only the switch that was nearest to the auxiliary phrase was considered for the analysis.

2.3 Results

2.3.1 Oral corpus

Out of the entire oral corpus, I extracted 845 tokens of sentences with the progressive structure. Within these tokens, 93 tokens constituted codeswitched sentences. Table 2-1 displays the way in which the codeswitches were distributed amongst these sentences.

Table 2-1: Distribution in the oral corpus of codeswitches in sentences including the progressive structure

<table>
<thead>
<tr>
<th>Codeswitch</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch immediately preceding the auxiliary</td>
<td>3</td>
<td>3.22%</td>
</tr>
<tr>
<td>Switch at the progressive auxiliary</td>
<td>7</td>
<td>7.53%</td>
</tr>
<tr>
<td>Switch at the present participle</td>
<td>7</td>
<td>7.53%</td>
</tr>
<tr>
<td>Switch immediately following the present participle</td>
<td>13</td>
<td>13.98%</td>
</tr>
<tr>
<td>Switch elsewhere (i.e., further away from the auxiliary phrase in either direction)</td>
<td>63</td>
<td>67.74%</td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>100%</td>
</tr>
</tbody>
</table>

The results show that out of the 93 tokens of codeswitched sentences that included the progressive structure, the majority of the switches (i.e., 63 of them) occurred at a considerable
distance from the auxiliary phrase. Of the remaining 30 tokens, 13 included a switch directly after the present participle, three immediately preceding the auxiliary, and seven right at the auxiliary. Moreover, seven instances of a switch between the Spanish auxiliary *estar* and an English present participle were found, representing 7.53% of the codeswitched tokens. The following sentences comprise the seven examples of switches at the present participle that were found.

(13) *o tal vez están developing un curso nuevo o algo así* (MIG, herring2)
   'or maybe (they) are developing a new course or something like that’

(14) *Paige yo no sabía que xxx estaban recording también* (SAR, herring10)
   'Paige, I did not know that xxx were recording too’

(15) *my oldest son went through a [/] a little bit once he got to the point que estaba dating* (JES, sastre9)
   'my oldest son went through a [/] a little bit once he got to the point when (he) was dating’

(16) *pero eso lo están (. testing ahora* (ISA, sastre11)
   'but that, (they) are (.) testing it now’

(17) *little [/] little marielitos will [//] van a estar ahí developing* (JAD, sastre7)
   'little [/] little marielitos will [//] will be there developing’

(18) *todos estamos con un calor and working* (CHL, herring1)
   'we are all hot and working’

(19) *porque cuando fueron a ver todo lo que pasó después del bombardeo cuando ya vieron a la gente que estaba prácticamente &e &e muerta o laying on [/] on [//] en [/] en el piso supuestamente estaban en ropa de dormir* (ISA, zeledon5)
   'because when (they) went to see all that happened after the bombing, when (they) saw the people who were practically &e &e dead or laying on [/] on [//] on the floor, supposedly (they) were in nightgowns’

The examples displayed here were produced by several different speakers. In examples (13), (14), (15), and (16) the Spanish auxiliary and the English present participle are immediately adjacent to each other, although in (16) there is a pause between the auxiliary and the participle, signaled by the symbol (.) In examples (17), (18), and (19), there is intervening material
–different adverbial expressions– between the auxiliary and the participle. It is worth noting that in example (17), the auxiliary *estar* appears in its infinitive form and not in its finite form because it is part of a more complex verbal structure (*van a estar ahí* developing ‘(they) will be there developing’). Nonetheless, the fact that the switch can be found whether the auxiliary is in its finite or nonfinite form is interesting. In addition, examples (18) and (19) display cases of ellipsis. Although there is no explicit auxiliary before the present participles *working* and *laying*, the presence of the conjunctions *and* and *o* ‘or’ indicates that the same Spanish auxiliary that was previously used in each case is also linked to the English present participles that follow the conjunctions.

Regarding sentences with the perfect structure, 375 tokens were extracted from the oral corpus. Of these 375 sentences, 28 tokens comprised codeswitched sentences. Table 2-2 shows the distribution of codeswitches in these sentences.

Table 2-2: Distribution in the oral corpus of codeswitches in sentences including the perfect structure

<table>
<thead>
<tr>
<th>Codeswitch</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch immediately preceding the auxiliary</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Switch at the perfect auxiliary</td>
<td>3</td>
<td>10.71%</td>
</tr>
<tr>
<td>Switch at the past participle</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Switch immediately following the past participle</td>
<td>6</td>
<td>21.43%</td>
</tr>
<tr>
<td>Switch elsewhere (i.e., further away from the auxiliary phrase in either direction)</td>
<td>19</td>
<td>67.86%</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2-2 shows that out of the 28 codeswitched sentences that included the perfect structure, in 19 of them, the switch occurred at a substantial distance from the auxiliary phrase. Moreover, in three of the sentences the switch took place at the auxiliary, and in six of them it appeared right after the participle. Crucially, no instances of switches between the Spanish *haber* auxiliary and the English past participle were found.
2.3.2 Written corpus

From the written corpus, 142 tokens of sentences with the progressive structure were extracted, and out of those, 106 tokens comprised codeswitched sentences. The distribution of these codeswitched sentences is shown in Table 2-3.

Table 2-3: Distribution in the written corpus of codeswitches in sentences including the progressive structure

<table>
<thead>
<tr>
<th>Codeswitch</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch immediately preceding the auxiliary</td>
<td>1</td>
<td>0.94%</td>
</tr>
<tr>
<td>Switch at the progressive auxiliary</td>
<td>8</td>
<td>7.55%</td>
</tr>
<tr>
<td>Switch at the present participle</td>
<td>8</td>
<td>7.55%</td>
</tr>
<tr>
<td>Switch immediately following the present participle</td>
<td>9</td>
<td>8.49%</td>
</tr>
<tr>
<td>Switch elsewhere (i.e., further away from the auxiliary phrase in either direction)</td>
<td>80</td>
<td>75.47%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>106</td>
<td>100%</td>
</tr>
</tbody>
</table>

Out of the total number of codeswitched sentences that included a progressive structure (i.e., 106), 80 sentences, representing approximately 75% of the codeswitched data, occurred at a distance from the auxiliary phrase. Of the remaining 26 tokens, nine included a switch right after the present participle, one included a switch at the word immediately preceding the auxiliary, and eight included a switch at the progressive auxiliary. Importantly, eight instances of switches between the auxiliary and the participle were found. The eight examples of codeswitches at the present participle are displayed below.6

(20) *Y hablando del Cervantes, te habrás enterado que los Catalanes están prohibiting the use of Spanish, decía un news que hasta los padres de school children are expected to speak Catalan, so there you are. (19/9/2008)*

‘And speaking of Cervantes, (you) must have found out that the Catalans are prohibiting the use of Spanish, a news [article] said that even the parents of school children are expected to speak Catalan, so there you are.’

---

6 The examples from the written corpus are reproduced word for word with their original spellings, which occasionally differ from standard Spanish and English.
(21) Y según dicen el Chief Minestra se ha cargao al Spanish-owned company que estaba fabricating los Waterport Terraces. (8/1/2010)
‘And, according to what (they) say, the Prime Minister has gotten rid of the Spanish-owned company that was fabricating the Waterport Terraces.’

(22) Dice que esta ticking. (15/1/2010)
‘(He) says (it) is ticking.’

(23) Bueno, hija, at least the weather parece que esta improving, porque vaya invierno que we are having, un poco de todo, frió, rain y un viento que te despeina. (5/2/2010)
‘Well, my child, at least the weather looks like (it) is improving, because this weather we are having is incredible, a little bit of everything, cold, rain and wind that messes up your hair.’

(24) Aquí están algunos saying nothing porque the Spaniards might take offence about nothing, don’t you agree Cloti dear? (29/2/2008)
‘Here are some [people] saying nothing because the Spaniards might take offense about nothing, don’t you agree Cloti dear?’

(25) A lo mejor they thought nobody would notice it porque estaba el mundo watching los Olympics. (15/8/2008)
‘Maybe they thought nobody would notice it because the world was watching the Olympics.’

(26) Well, it’s over 200 pages long – y está todo el mundo financiero making copies of it, not to mention los lawyers y otros curiosos. (12/11/2004)
‘Well, it’s over 200 pages long – and the entire financial world is making copies of it, not to mention the lawyers and other curious [people].’

(27) Hasta BenTata esta dando la lata, saying que it is nonsensical to open el Instituto Cervantes con lo de palos que nos estan dando. (26/11/2010)
‘Even BenTata is bothering (us), saying that it is nonsensical to open the Cervantes Institute with the beating (they) are giving us.’

In examples (20)-(23), the Spanish auxiliary and the English present participle are immediately adjacent to each other. In examples (24)-(26), there is intervening material between the auxiliary and the participle. In these cases, the intervening material always constitutes the subject of the clause, either a pronoun or a full noun phrase. Example (27) comprises another case of ellipsis, in which the Spanish auxiliary is linked not only to the first present participle (dando), but also to the second present participle that is located after the comma (saying).
Concerning sentences that included the perfect structure, 183 tokens were extracted from the written corpus and 150 of them constituted codeswitched sentences. Table 2-4 exhibits the distribution of codeswitches in these sentences.

Table 2-4: Distribution in the written corpus of codeswitches in sentences including the perfect structure

<table>
<thead>
<tr>
<th>Codeswitch</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch immediately preceding the auxiliary</td>
<td>1</td>
<td>0.66%</td>
</tr>
<tr>
<td>Switch at the perfect auxiliary</td>
<td>15</td>
<td>10%</td>
</tr>
<tr>
<td>Switch at the past participle</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Switch immediately following the past participle</td>
<td>28</td>
<td>18.67%</td>
</tr>
<tr>
<td>Switch elsewhere (i.e., further away from the auxiliary phrase in either direction)</td>
<td>106</td>
<td>70.67%</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>100%</td>
</tr>
</tbody>
</table>

Here, out of the 150 codeswitched sentences with the perfect structure, approximately 70% of the tokens included switches that occurred at a distance from the auxiliary phrase. From the remaining 44 tokens, 28 included a switch right after the past participle, one included a switch on the word that preceded the auxiliary, and 15 included a switch at the perfect auxiliary. Finally, no instances of a switch between the haber auxiliary and the past participle were found.

### 2.4 Discussion

This corpus study examined the appearance of codeswitches involving the progressive and perfect auxiliary phrases in oral and written corpora. Overall, there was much more codeswitching present in the written corpus than in the oral corpus, at least within sentences that included the progressive and perfect structures. In the written corpus, the codeswitched tokens represented approximately 82% of the total amount of extracted sentences of progressive and perfect structures, while in the oral corpus, they only represented 11%. When only the codeswitched data are considered, it becomes evident that the amount of codeswitched sentences
that included the progressive structure was similar across both corpora (93 tokens in the oral
corpus and 106 in the written corpus). There were larger differences between both corpora
regarding the amount of codeswitched sentences that included the perfect structure. The oral
corpus included only 28 tokens whereas the written corpus included 150 tokens of these
sentences. The large difference between the two corpora regarding the appearance of
codeswitched sentences with the perfect structure may simply have to do with the topics
discussed in each corpus; the Calentita editorial columns tend to refer to events that have taken
place in the near past while the oral conversations might include less narratives referring to past
events and more referring to the present or future. The discrepancy may also be explained by
dialectal differences. While the perfect structure is used frequently in Peninsular Spanish
dialects, such as the one used in Gibraltar, it is less common in Latin American Spanish dialects,
particularly those spoken by most of the individuals in the Miami corpus. In these dialects, the
simple past tense is used more often.

Despite this difference in overall occurrence of progressive and perfect structures, the
oral and written corpora are remarkably similar with respect to the distribution of the different
switch locations examined here. First, in sentences that included both the progressive and the
perfect structure, the majority of switches occurred at a considerable distance from the auxiliary
phrase, either before or after it. In Tables 2-1 to 2-4, these switches accounted for 67% to 75% of
the codeswitched data. The next most common switch site in both corpora for sentences that
included progressive and perfect structures was immediately after the present participle or past
participle, respectively. The amount of data for which these switches accounted ranged from 8% to 21%. Finally, switches on the word that immediately preceded the auxiliary were almost
nonexistent, never accounting for more than 3% of the codeswitched data.

Clear differences between the codeswitched sentences that included the progressive and
perfect structures arise when it comes to switches directly involving the auxiliary phrase. In the
oral corpus, seven sentences included a switch at the progressive auxiliary and another seven included a switch at the present participle, each of these representing 7.53% of the codeswitched data. Therefore, a switch was as likely to occur at the auxiliary as it was to occur at the present participle. However, turning to sentences with the perfect structure, three (10.71%) included a switch at the perfect auxiliary and none included a switch at the past participle. In the written corpus, a very similar picture emerged. In this case, eight sentences included a switch at the progressive auxiliary and an additional eight sentences included a switch at the present participle, each of these representing 7.55% of the codeswitched data. Thus, in the written corpus it also seems that a switch at the progressive auxiliary and a switch at the present participle are equally probable. These percentages of occurrence are strikingly similar between the two corpora.

Regarding written sentences that include the perfect structure, in 15 of them the switch occurred at the perfect auxiliary, representing 10% of the codeswitched data, and there were no occurrences of switches at the past participle. Once again, these results closely resemble those in the oral corpus.

It is also worth noting that the particular examples of switches at the present participle were similar across both corpora. In oral and written form, there were cases in which the *estar* auxiliary and the English present participle were adjacent to each other and other cases in which there was intervening linguistic material. In addition, in both corpora the codeswitched progressive structure appeared with varied grammatical subjects (i.e., noun phrases, pronouns, omitted subjects) and the progressive auxiliary appeared in different tenses (i.e., present, past). The codeswitches that are examined in this dissertation are specifically switches between Spanish auxiliaries and English participles. Although switches in the other direction, that is, switches between English auxiliaries and Spanish participles were not considered in this corpus study, two sentences including this type of switch were found in the corpora examined here. Interestingly,
one example was found in each corpus and they were both examples of switches involving the progressive structure. They are displayed below.

(28) no he [/] he was jodiendo (EMI, sastre4)
    ‘no, he [/] he was joking’

(29) And there he is in Madrid paseándose por la calle de Alcalá, y como no tenga ciudadó termina bailando un chotis. (Calentita, 10/7/2009)
    ‘And there he is in Madrid walking around Alcalá Street, and, if (he) is not careful, (he) will end up dancing a chotis.’

Taken together, these results reveal that estar+English participle switches are more frequently found in natural oral and written production corpora than haber+English participle switches. According to these data, Spanish-English bilinguals seem to be equally likely to switch at the progressive auxiliary or at the present participle when using the progressive structure. However, when using the perfect structure, they seem to avoid switching at the participle. Instead, they prefer to switch at the auxiliary or right after the participle. The findings from this corpus study coincide with previously published reports, which also exhibit the more frequent appearance of estar+English participle switches, when compared to haber+English participle switches, in spontaneous production data (Lipski, 1978, 1985; Pfaff, 1979; Poplack, 1980).
Chapter 3

Comprehension costs associated with auxiliary phrase switches

The results of the corpus study described in Chapter 2, coupled with examples and reports on the distribution of codeswitches in Spanish-English naturalistic production, demonstrate the discernible frequency with which estar+English participle and haber+English participle switches occur. Namely, switches involving the progressive structure are more often produced than switches involving the perfect structure. It is important to note that, although the progressive and perfect structures display different tense and aspect features and, thus, inherently express different temporal meanings, whether a codeswitch is produced at the participle verb or at the auxiliary does not alter the meaning of the utterance. In other words, codeswitching at either one of those points in the phrase does not add to or subtract from the sense of the utterance.

Turning back to the examples mentioned in Table 1-1, ‘mi marido está working on his Master’s’ (‘my husband is working on his Master’s,’ example taken from Lipski, 1978, p. 252) has the exact same meaning as ‘mi marido is working on his Master’s.’ Similarly, ‘Yo creo que apenas se había washed out’ (‘I think that it had just washed out,’ example taken from Pfaff, 1979, p. 300) means the same thing as ‘Yo creo que apenas la cosa had washed out’ (‘I think that the thing had just washed out’). Therefore, the frequency with which these codeswitches do or do not occur is not related to a change in the meaning or sense of the auxiliary phrase. Regardless of the reasons behind the difference in the frequency of occurrence of these two auxiliary phrase switches, for now the crucial point is that this difference exists.
3.1 Research questions

Based on these production findings, a study was conducted to see whether production patterns were reflected in comprehension costs. We now know that *estar*+English participle switches are more frequent than *haber*+English participle switches; the question remains whether the former type of switch is easier to process or comprehend than the latter. The study was designed to address this question. In other words, the first goal of the study was to see whether codeswitches that are more frequently found in production corpora (i.e., *estar*+English participle switches) are easier to process than codeswitches that are relatively less frequent (i.e., *haber*+English participle switches). The second goal of the study was to see whether the age at which an individual acquires the second language (L2) influences the way these two codeswitches are processed. Childhood bilinguals and late bilinguals were recruited because previous work on codeswitching that examined the oral exchanges of Spanish-English speakers living in bilingual communities have found that age of L2 acquisition can influence the type of codeswitches that they produce. Poplack (1980) showed that individuals who were born and raised in a bilingual environment and who were exposed to codeswitching from early on, produced comparatively more intrasentential switches than bilinguals whose acquisition of the L2 (English) took place in adulthood. In other words, the nature of the codeswitches produced by the two types of bilinguals was starkly different. Whereas childhood bilinguals produced more alternating codeswitches, late (i.e., Spanish-dominant) bilinguals produced more tag switches. If age of L2 acquisition affects the types of switches that bilinguals produce, there may also be a difference in the way in which early acquirers of both languages and late acquirers of the L2 process intrasentential switches, like the ones examined here. Finally, the third goal was to see if and how different tasks affected the way these two groups of bilinguals processed *estar*+English participle switches and *haber*+English participle switches. Two tasks were compared: a
comprehension task and an acceptability judgment task. The first task is widely used in psycholinguistic studies that examine linguistic processing whereas the second task has been the preferred task used in codeswitching studies that aim to model bilingual competence. To summarize, the three research questions addressed in this study are:

1) Are more frequently produced types of codeswitches easier to process by the comprehension system than less frequent types of codeswitches?
2) Does age of second language acquisition play a role in the comprehension of these types of switches?
3) Do different tasks influence processing patterns during the comprehension of these codeswitches?

In order to address these questions, the study used the eye-tracking technique to measure participants’ fixation durations as they read the codeswitched sentences that appeared on a computer screen. Longer fixation durations were considered an index of increased comprehension costs.

3.2 Method

3.2.1 Participants

The data for this study was collected between January 2009 and December 2010 in State College, Pennsylvania. Forty-seven participants took part in this study. Eleven of these participants were excluded from the study for several reasons. Three participants were excluded because of technical problems with the eye-tracking equipment; three additional participants were eliminated because they were native English speakers who had learned Spanish as adults in a classroom setting; two participants reported that they did not codeswitch at all; two additional
participants were exposed to a language other than English or Spanish in their home; and one participant reported living in the United States (U.S.) for only one month. The remaining 36 participants were divided into two groups: a group of 18 early bilinguals and a group of 18 late bilinguals. Participants were placed into one of these two groups based on a series of characteristics and scores obtained from several measures of language proficiency described below.

3.2.1.1 Language History Questionnaire

In an online Language History Questionnaire (LHQ), participants provided self-ratings of their English and Spanish proficiency across reading and writing production as well as speaking and listening comprehension. They also answered open-ended and multiple-choice questions about their history with both languages, their language learning experiences, and their daily exposure to and use of both languages. Several examples of the questions asked in the LHQ are included in Appendix A.

3.2.1.2 Boston Naming Vocabulary Test

As a measure of lexical access, vocabulary size, and naming performance, participants completed the Boston Naming Vocabulary Test (BNT, Kaplan, Goodglass, & Weintraub, 1983) in English and Spanish. The BNT contains 60 outline drawings of objects and animals. The images were divided into two language blocks (and English block and a Spanish block) of 30 images each. Participants completed the BNT in their self-reported dominant language first, and then in the other language. In each language block, participants were asked to name the images as quickly and as accurately as possible. The drawings were presented in order of increasing
difficulty, starting with easy, high-frequency words, such as flower and cama ‘bed’ and concluding with more difficult, low-frequency words like protractor and yunta ‘yoke.’ Appendix B includes examples of the outline drawings used in each language block. Participants’ responses to the BNT were digitally recorded and these recordings were then revised in order to score the test. Participants received a score of 1 for every correctly named image and 0 for every incorrectly named image or for any unnamed image.

3.2.1.3 Grammatical competency tests

Participants also completed the grammar sections of the Michigan English Language Institute College Entrance Test (MELICET) and the Advanced Test of the Diplomas de Español como Lengua Extranjera (DELE, ‘Diplomas of Spanish as a Foreign Language’). The MELICET is an advanced level English language test created by the University of Michigan English Language Institute (http://www.michigan-proficiency-exams.com/melicet.html) to examine ability in different English language areas. It is primarily used to test nonnative speakers of English by educational institutions as an admissions or placement test. The DELE is a standardized test of Spanish issued by the Ministry of Education, Culture, and Sport of Spain, which assesses proficiency in Spanish at seven levels (http://diplomas.cervantes.es/en). The test administered here was the Nivel Superior C2, the highest level of accreditation. Each grammar test contains 50 multiple-choice items, which evaluate grammar, vocabulary, and reading competence in isolated sentences as well as longer stretches of discourse. Appendix C displays examples of the items included in the MELICET and the DELE. Participants received one point for each correct answer and no points for incorrect answers.

All participants were undergraduate or graduate students at The Pennsylvania State University. The early bilinguals were born and raised in the U.S., with the exception of 5
participants, who were born in Spanish-speaking countries and moved to the U.S. during childhood or adolescence. The late bilinguals were born in Spanish-speaking countries and moved to the U.S. during adolescence or adulthood, with the exception of one participant who was born and raised in the U.S. and three who spent several childhood years in the U.S. In addition, all participants were proficient in both Spanish and English. They reported regular use of and exposure to both languages in oral and written mode. Importantly, they also codeswitched occasionally or frequently with other bilinguals. More specific participant characteristics are displayed in Table 3-1.

Table 3-1: Characteristics of the State College participant groups

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Early bilinguals n = 18</th>
<th>Late bilinguals n = 18</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Range)</td>
<td>Mean (Range)</td>
</tr>
<tr>
<td>Age</td>
<td>21 (18-37)</td>
<td>22 (18-31)</td>
</tr>
<tr>
<td>Age of arrival in U.S.</td>
<td>3 (0-13)</td>
<td>16 (0-29)</td>
</tr>
<tr>
<td>Years spent living in U.S.</td>
<td>19 (8-37)</td>
<td>5 (1-18)</td>
</tr>
<tr>
<td>Age of reported achieved fluency in English</td>
<td>7 (4-14)</td>
<td>14 (6-29)</td>
</tr>
<tr>
<td>Age of reported achieved fluency in Spanish</td>
<td>7 (2-16)</td>
<td>4 (2-8)</td>
</tr>
<tr>
<td>Difference in years between reported achieved fluency in L1 and L2</td>
<td>5 (0-13)</td>
<td>10 (2-26)</td>
</tr>
<tr>
<td>Self-ratings for English proficiency (/10)</td>
<td>9.3 (7.5-10)</td>
<td>8.5 (5.75-9.75)</td>
</tr>
<tr>
<td>Self-ratings for Spanish proficiency (/10)</td>
<td>8.2 (6-10)</td>
<td>9.4 (7.5-10)</td>
</tr>
<tr>
<td>BNT score for English (/30)</td>
<td>21 (15-29)</td>
<td>17 (8-23)</td>
</tr>
<tr>
<td>BNT score for Spanish (/30)</td>
<td>13 (3-24)</td>
<td>21 (17-25)</td>
</tr>
<tr>
<td>MELICET score (/50)</td>
<td>45 (36-49)</td>
<td>40 (31-47)</td>
</tr>
<tr>
<td>DELE score (/50)</td>
<td>33 (11-43)</td>
<td>42 (33-48)</td>
</tr>
</tbody>
</table>

Table 3-1 shows that, in general, the group of late bilinguals arrived in the U.S. later than the group of early bilinguals. The late bilinguals have also been living in the U.S. for less time than the early bilinguals. Moreover, the early bilinguals achieved fluency in both Spanish and English at similar ages whereas the late bilinguals became fluent in their L2 (English) at a later age. Despite this difference, these four participants were still included in the group of late bilinguals because they fit in better with this group based on other measures of language proficiency, such as language self-ratings and scores on vocabulary naming tests and grammar tests. Like the rest of the late bilinguals, these four participants presented higher self-ratings and test scores in Spanish.
age. Therefore, the late bilinguals display a larger difference between achieved fluency in their first and second language. Finally, the early bilinguals have higher self-ratings of English proficiency as well as higher English vocabulary naming and grammar scores. The opposite is true for the group of late bilinguals, who present higher Spanish proficiency self-ratings and test scores. The results of several independent-samples t tests showed significant differences between both groups of participants for the proficiency measures displayed in Table 3-1 (self-ratings for English proficiency: \( t(34) = 2.81, p = .008 \); self-ratings for Spanish proficiency: \( t(34) = 3.82, p = .001 \); BNT score for English: \( t(34) = 3.06, p = .004 \); BNT score for Spanish: \( t(34) = 5.71, p < .001 \); MELICET score: \( t(34) = 3.14, p = .004 \); DELE score: \( t(34) = 4.25, p < .001 \).

3.2.2 Materials and Design

The experimental stimuli comprised 96 item sets (Appendix D). Each item set consisted of four different versions of the same sentence, corresponding to four experimental conditions. Conditions 1 and 2 were codeswitched conditions with the progressive structure. In Condition 1, the switch occurred at a phrasal boundary, that is, right at the auxiliary. Condition 2 contained a switch within the auxiliary phrase (between the Spanish auxiliary estar and the English present participle). Conditions 3 and 4 were analogous to Conditions 1 and 2, but involved the perfect structure instead. In each experimental sentence, the critical region under examination was part of an embedded phrase to ensure its appearance in the middle of the sentence and, thus, in the middle of the computer screen. Table 3-2 displays an example item set, in which the critical region is underlined.
The 96 item sets were divided into six reading files, each of which included 32 experimental codeswitched sentences (eight sentences for each condition). Participants were never exposed to the same sentence in more than one condition. In addition to the experimental sentences, 32 codeswitched sentences were added as fillers. Both the experimental sentences and the fillers were controlled as much as possible to ensure that extraneous factors were not responsible for the predicted pattern of results. First, the experimental stimuli were controlled for word length (mean = 13, range = 11-14). In addition, given that the main verb in the sentence was always followed by a sentential complement that included the critical region under examination, only main verbs that had a sentential complement bias or verbs that were equi-biased for sentential complements or direct objects were used, in order to avoid processing difficulties due to main verb subcategorization preferences. The grammatical subjects of the verbs in the critical region had a mean length of 10 characters (range 5-14) and a mean lexical frequency of 27.64 (range 1-162). The lexical frequencies for Spanish words are from the Alameda and Cuetos (1995) corpus, which is based on 2 million words, and they were obtained through the Normas e índices de interés en Psicología Experimental (NIPE, ‘Norms and Indices of interest in Experimental Psychology’) website (Díez, Fernández, & Alonso, 2006). Moreover, these grammatical subjects were always a cognate noun in Spanish and English to maximize

<table>
<thead>
<tr>
<th>Condition</th>
<th>Sample sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-switch at auxiliary</td>
<td>El director confirmó que los actores are rehearsing their lines for the movie.</td>
</tr>
<tr>
<td>(2) Progressive-switch at participle</td>
<td>El director confirmó que los actores están rehearsing their lines for the movie.</td>
</tr>
<tr>
<td>(3) Perfect-switch at auxiliary</td>
<td>El director confirmó que los actores have rehearsed their lines for the movie.</td>
</tr>
<tr>
<td>(4) Perfect-switch at participle</td>
<td>El director confirmó que los actores han rehearsed their lines for the movie.</td>
</tr>
</tbody>
</table>

Translation: ‘The director confirmed that the actors are rehearsing/have rehearsed their lines for the movie.’
crosslinguistic lexical activation. The present participles and the past participles were only from regular-ending verbs in order to keep the spelling of the participles as uniform as possible. The mean length of the present participles in the critical region was nine characters (range 6-11) and that of the past participles was eight characters (range 5-10). Mean lexical frequency of the present participles was 16.03 (range 1-76) and that of the past participles was 51.68 (range 3-401). The lexical frequencies for English words are from the Kučera and Francis (1967) corpus, which is based on 1 million words, and they were obtained through The English Lexicon Project website (Balota et al., 2007).

The fillers were similar to the experimental items in terms of overall length, but differed from them regarding the syntactic structures and the codeswitch types included. Three examples of the fillers are provided below.

(30) Switch between the verb and the direct object
Laura estaba limpiando the kitchen before going out with her friends.
‘Laura was cleaning the kitchen before going out with her friends.’

(31) Switch between the definite article and the noun
Tomás y su esposa ya habían visto el movie that their friends had recommended.
‘Thomas and his wife had already seen the movie that their friends had recommended.’

(32) Switch between clauses
Como la maestra ha sospechado, the students have not studied for the exam.
‘As the teacher has suspected, the students have not studied for the exam.’

Additional care was taken with respect to task demands. For the comprehension task, half of the sentences were followed by questions whose response was positive, while the other half was followed by questions whose response was negative. Moreover, half of the questions referred to the beginning of the sentence and the other half referred to the end of the sentence. For the acceptability judgment task, half of the filler sentences included a grammatical error (e.g., lack of subject-verb agreement, lack of gender or number agreement, incorrect use of verb tense or
mood) at the beginning of the sentence and the other half included the grammatical error towards the end of the sentence. Two examples of the ungrammatical fillers are displayed below.

(33) Gender agreement error at the beginning of the sentence
*El enfermera había traído the soup for my sister while she was asleep.
‘The masculine nurse feminine had brought the soup for my sister while she was asleep.’

(34) Number agreement error towards the end of the sentence
*Los ingenieros estuvieron dibujando the plans for all the house on the corner.
‘The engineers were drawing the plans for all the house on the corner.’

The sentences within each of the 6 reading files were pseudo-randomly interleaved to avoid order effects. Participants completed the comprehension task and the acceptability judgment task in two separate blocks. Block presentation order was counterbalanced across participants.

3.2.3 Procedure

Before coming to the laboratory to complete the eye-tracking experiment, participants filled out the online LHQ. Once in the laboratory, stimuli for the eye-tracking experiment were presented on a color monitor using an EyeLink 1000 desktop-mounted eye-tracker, interfaced with an IBM-compatible PC. With this system, participants sat in front of the computer screen at a distance of approximately 75 centimeters and used a chin rest and a forehead pad to minimize head movement. Eye movements were recorded with a camera and an infrared illuminator, located at the bottom of the computer monitor. Monocular tracking of the right pupil and cornea was performed at a sampling rate of 1000 hertz. The eye-tracker was calibrated and validated for each participant at the start of each experimental block and after each break to calculate overall equipment accuracy. Each sentence was displayed across one line in the middle of the computer screen, in Consolas font type, size 14. Occasionally, in the case of longer sentences, several
words were displayed on a second line. Crucially, the critical region of the experimental sentences always appeared in the middle of the first line.

Participants completed the experiment in one session that lasted approximately one hour and 30 minutes. During eye-tracking, participants were instructed to read each sentence silently at their own pace. In the comprehension block, after reading the sentence, participants were asked to answer a comprehension question related to the content of the sentence. The questions were answered with either “yes” or “no,” by pressing one of two buttons on a game pad. During the acceptability judgment block, participants were asked to use the same two buttons on the game pad to answer “yes” if they thought that the codeswitched sentence sounded like the types of codeswitches that they were used to hearing and using, and “no” if they thought that the codeswitch did not sound right. In addition to the eye-tracking experiment proper, participants completed the tasks described above in section 3.2.1, which were used to assess language proficiency (i.e., vocabulary naming tasks and grammar tests). These tasks were completed between the eye-tracking blocks to avoid participant fatigue. Participants were compensated with $15 for their participation.

3.3 Predictions

It is important to note that the data obtained from the eye-tracker are different measures of reading speed. Reading speed is taken as a reflection of processing costs. Thus, results of quickly read words or phrases are interpreted as evidence of easy processing whereas results of slowly read words or phrases are taken as proof of costly processing. Based on the three research questions, the following predictions were formulated. A comparison between sentences with switches at a phrasal boundary and those with switches within the auxiliary phrase should reveal differential costs associated with distinct types of codeswitches. Therefore, regarding the first
research question, if frequently produced codeswitch types are processed with more ease than less frequent codeswitch types, then *haber*+English participle switches (Condition 4; e.g., *los actores han* rehearsed ‘the actors have rehearsed’) should produce longer reading times than their corresponding phrasal boundary switch (Condition 3; e.g., *los actores* have rehearsed). This reading time difference would reflect the fact that phrasal boundary switches are common in naturalistic bilingual expression, while *haber*+English participle switches are virtually nonexistent. However, there should not be a reading time difference between *estar*+English participle switches (Condition 2; e.g., *los actores están* rehearsing ‘the actors are rehearsing’) and their corresponding phrasal boundary switches (Condition 1; e.g., *los actores* are rehearsing) because both types of codeswitches are found in spoken and written corpora with approximately the same frequency of occurrence.

Concerning the second research question, if the age in which an L2 is acquired influences comprehension costs of codeswitched sentences, then early bilinguals and late bilinguals should display different reading times on the critical region under examination. Specifically, because of their increased experience with the L2 and, potentially, with codeswitching, the early bilinguals’ reading times should better reflect the codeswitching patterns found in naturalistic production corpora, compared to the late bilinguals. Turning to the third research question, if different tasks influence patterns of sensitivity during the comprehension of these codeswitches, the comprehension task and the acceptability judgment task should produce distinct reading time results on the critical region of the sentences.

3.4 Results

The critical region for which reading measures were extracted was the participle (present participle in the case of the progressive structure or past participle in the case of the perfect
structure) in the experimental sentences. The participle was selected as the critical region because it is at that point in the sentence where the participants have processed the complete auxiliary phrase. It is also at that point where all codeswitches, both the codeswitches at the auxiliary and the codeswitches at the participle, have occurred. Therefore, any processing costs encountered at that point include any potential switch costs of the entire (codeswitch or unswitched) auxiliary phrase. Three eye-tracking measures were extracted for analysis: gaze duration, regression path time, and total time. Gaze duration refers to the sum of all fixation durations in the critical region (i.e., the participle) from first entering it until leaving it (Rayner, 1998, p. 377). Regression path time is the sum of all fixation durations in the critical region from first entering it until leaving it to proceed further in the sentence by moving towards the right of the critical region. This reading measure includes any leftward movements to words that precede the critical region (Rayner, 1998, p. 377). Total time represents the sum of all fixation durations in the critical region, including all regressive fixation durations to it (Rayner & Duffy, 1986, p. 196). These three eye-tracking measures were chosen because they reflect both early and late comprehension processes. Gaze duration and regression path time are generally considered to be sensitive to early processes in the comprehension of a sentence, such as lexical characteristics, syntactic parsing, and early integration of information. Total time is assumed to be sensitive to the later processes involved in the comprehension of sentences, such as re-analysis and discourse integration (Clifton, Staub, & Rayner, 2007; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

Two measures of early processing, instead of just one, are extracted in order to account for differences in participants’ reading practices. In eye-tracking studies, several measures are usually extracted because it is difficult to know exactly at what point during reading each participant processes the information. Participants use different strategies when they encounter difficulty during reading. For example, some participants pause on a word or phrase as soon as they encounter difficulty, which leads to longer gaze durations. Others tend to produce shorter
initial fixation durations, but then return to the areas of the sentence that were difficult to process and re-read them. This way of reading produces shorter gaze durations, but longer regression path reading times. Because of these differences between participants, it is important to extract more than one measure of early processing (Carreiras & Clifton, 1999). The results that correspond to the comprehension task will be addressed first, followed by the results for the acceptability judgment task.

### 3.4.1 Results for the comprehension task

Table 3-3 displays the proportions of correct answers to the comprehension questions by condition and participant group. The results show that, although the proportions for the early bilinguals are slightly higher than those for the late bilinguals, overall both groups of participants answered most comprehension questions correctly, demonstrating that they were paying attention to the task and that they understood the content of the sentences.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent of correct answers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early bilinguals</td>
</tr>
<tr>
<td>(1) Progressive-Switch at auxiliary</td>
<td>93.06%</td>
</tr>
<tr>
<td>(2) Progressive-Switch at participle</td>
<td>90.97%</td>
</tr>
<tr>
<td>(3) Perfect-Switch at auxiliary</td>
<td>90.97%</td>
</tr>
<tr>
<td>(4) Perfect-Switch at participle</td>
<td>90.28%</td>
</tr>
</tbody>
</table>
3.4.1.1 Early bilinguals

Table 3-4 presents the mean gaze duration, regression path time, and total time by condition for the group of early bilinguals. The standard deviation for each mean is provided in parentheses.

Table 3-4: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the early bilinguals during the comprehension task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-</td>
<td>306.59</td>
<td>399.81</td>
<td>478.66</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>(SD = 113.51)</td>
<td>(SD = 233.44)</td>
<td>(SD = 176.81)</td>
</tr>
<tr>
<td>(2) Progressive-</td>
<td>337.74</td>
<td>409.07</td>
<td>523.80</td>
</tr>
<tr>
<td>Switch at participle</td>
<td>(SD = 55.92)</td>
<td>(SD = 98.36)</td>
<td>(SD = 124.94)</td>
</tr>
<tr>
<td>(3) Perfect-</td>
<td>298.67</td>
<td>386.15</td>
<td>500.77</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>(SD = 71.95)</td>
<td>(SD = 132.48)</td>
<td>(SD = 151.74)</td>
</tr>
<tr>
<td>(4) Perfect-</td>
<td>423.05</td>
<td>516.85</td>
<td>733.40</td>
</tr>
<tr>
<td>Switch at participle</td>
<td>(SD = 115.30)</td>
<td>(SD = 187.60)</td>
<td>(SD = 278.89)</td>
</tr>
</tbody>
</table>

A two-way repeated measures analysis of variance (ANOVA) was conducted to evaluate the effect of auxiliary type and switch site on the three extracted reading measures (gaze duration, regression path time, and total time). Auxiliary type (progressive versus perfect) and switch site (switch at the auxiliary versus switch at the participle) were the within-subjects factors. Separate analyses were conducted for each task and for each participant group.

Gaze duration. For gaze duration, the results yielded a main effect of auxiliary type, $F(1,17) = 5.21, p = .036$, a main effect of switch site, $F(1,17) = 39.74, p < .001$, and a significant by-participant interaction of auxiliary and switch site, $F(1,17) = 4.76, p = .043$. Paired-samples t tests were conducted to follow up the significant interaction. They indicated significant mean differences between Conditions 3 and 4, $t(17) = 5.31, p < .001$. However, no significant mean differences were found between Conditions 1 and 2, $t(17) = 1.21, p = .244$. Therefore, during the early reading measure of gaze duration, it took early bilinguals significantly longer to read *haber*+English participle switches than perfect structures in which the switch occurred at the
auxiliary. However, they read *estar*+English participle switches and progressive structures in which the switch occurred at the auxiliary at a similar speed.

**Regression path time.** For regression path time, the ANOVA did not exhibit a significant main effect of auxiliary type, $F(1,17) = 2.06, p = .170$, but it did yield a main effect of switch site, $F(1,17) = 5.65, p = .029$. A marginally significant by-participant interaction of auxiliary type and switch site was found, $F(1,17) = 4.32, p = .053$. Subsequent pairwise contrasts indicated significant mean differences between Conditions 3 and 4, $t(17) = 3.93, p = .001$, but not between Conditions 1 and 2, $t(17) = .19, p = .850$. These results demonstrate that during this second measure of early processing, the early bilinguals read the two codeswitched conditions with the progressive structure at a similar speed. Conversely, regarding the two codeswitched conditions with the perfect structure, switches at the participle were read at a significantly slower rate than switches at the auxiliary.

**Total time.** For total time, the results of the repeated measures ANOVA yielded a main effect of auxiliary type, $F(1,17) = 8.29, p = .01$, and of switch site, $F(1,17) = 15.19, p = .001$. Moreover, the results displayed a significant by-participant interaction of auxiliary type and switch site, $F(1,17) = 13.49, p = .002$. Once more, the follow-up paired-samples $t$ tests indicated significant mean differences between Conditions 3 and 4, $t(17) = 4.97, p < .001$, but not between Conditions 1 and 2, $t(17) = 1.11, p = .282$. When compared to their corresponding baseline conditions in which switches occurred at the auxiliary, *haber*+English participle switches were read significantly more slowly than *estar*+English participle switches.
3.4.1.2 Late bilinguals

The mean gaze duration, regression path time, and total time by condition for the group of late bilinguals is displayed in Table 3-5. Standard deviations for the means are shown in parentheses.

Table 3-5: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the late bilinguals during the comprehension task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-Switch at auxiliary</td>
<td>343.58 (SD = 116.53)</td>
<td>531.04 (SD = 227.19)</td>
<td>722.92 (SD = 379.29)</td>
</tr>
<tr>
<td>(2) Progressive-Switch at participle</td>
<td>364.27 (SD = 122.06)</td>
<td>521.87 (SD = 258.36)</td>
<td>685.65 (SD = 254.33)</td>
</tr>
<tr>
<td>(3) Perfect-Switch at auxiliary</td>
<td>316.72 (SD = 94.86)</td>
<td>469.88 (SD = 312.39)</td>
<td>668.52 (SD = 320.68)</td>
</tr>
<tr>
<td>(4) Perfect-Switch at participle</td>
<td>383.50 (SD = 131.24)</td>
<td>623.14 (SD = 303.93)</td>
<td>886.51 (SD = 276.23)</td>
</tr>
</tbody>
</table>

**Gaze duration.** For gaze duration, the results of the late bilinguals indicated no main effect of auxiliary type, $F(1,17) = .24, p = .877$, no main effect of switch site, $F(1,17) = .24, p = .628$, and no by-participant interaction of auxiliary type and switch site, $F(1,17) = 1.30, p = .270$.

Therefore, the group of late bilinguals displayed no significant participle reading time differences among any of the four experimental conditions.

**Regression path time.** The regression path time results did not display a main effect of auxiliary type, $F(1,17) = .24, p = .628$, but they did display a main effect of switch site, $F(1,17) = 11.80, p = .003$. There was also a significant by-participant interaction of auxiliary type and switch site, $F(1,17) = 4.78, p = .043$. Subsequent pairwise contrasts indicated significant mean differences between Conditions 3 and 4, $t(17) = 4.44, p < .001$, but not between Conditions 1 and 2, $t(17) = 1.85, p = .855$. For this reading measure, the late bilinguals read perfect structures in which the switch occurred at the auxiliary significantly more quickly than perfect structures in which the
switch occurred at the participle. They did not, however, read the two types of switches involving the progressive structure at significantly different reading speeds.

**Total time.** For total time, the results of the ANOVA indicated no main effect of auxiliary type, \(F(1,17) = 2.49, p = .133\), as well as no main effect of switch site, \(F(1,17) = 3.63, p = .074\). There was, however, a significant by-participant interaction of auxiliary type and switch site, \(F(1,17) = 9.29, p = .007\). Follow-up paired-samples \(t\) tests exhibited yet again significant mean differences between Conditions 3 and 4, \(t(17) = 4.02, p = .001\), but not between Conditions 1 and 2, \(t(17) = .52, p = .607\). In other words, in sentences with the perfect structure, late bilinguals read the participles significantly more slowly when the switch occurred at the participle than when it occurred at the auxiliary. Nonetheless, they read the participles at a similar speed in both types of experimental sentences with the progressive structure.

### 3.4.2 Results for the acceptability judgment task

Table 3-6 displays the proportions of acceptable judgments of the codeswitched sentences by condition and participant group.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent of acceptable judgments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early bilinguals</td>
</tr>
<tr>
<td>(1) Progressive-Switch at auxiliary</td>
<td>95.14%</td>
</tr>
<tr>
<td>(2) Progressive-Switch at participle</td>
<td>91.67%</td>
</tr>
<tr>
<td>(3) Perfect-Switch at auxiliary</td>
<td>95.14%</td>
</tr>
<tr>
<td>(4) Perfect-Switch at participle</td>
<td>79.86%</td>
</tr>
</tbody>
</table>

Table 3-6 shows that overall both conditions in which the switch occurred at the auxiliary (Conditions 1 and 3) were judged acceptable by both groups of participants. Regarding the conditions in which the switch occurred at the participle, both early and late bilinguals tended to
give sentences with *estar*+English participle switches (Condition 2) high judgments of acceptability. Sentences with *haber*+English participle switches (Condition 4), however, received lower proportions of acceptable judgments, particularly by the late bilinguals.

### 3.4.2.1 Early bilinguals

Table 3-7 presents the means and standard deviations for gaze duration, regression path time, and total time by condition for the group of early bilinguals.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-switch at auxiliary</td>
<td>333.71 (SD = 99.42)</td>
<td>437.58 (SD = 180.19)</td>
<td>513.82 (SD = 196.29)</td>
</tr>
<tr>
<td>(2) Progressive-switch at participle</td>
<td>416.80 (SD = 102.52)</td>
<td>543.20 (SD = 172.47)</td>
<td>747.22 (SD = 235.50)</td>
</tr>
<tr>
<td>(3) Perfect-switch at auxiliary</td>
<td>323.51 (SD = 81.60)</td>
<td>446.00 (SD = 147.67)</td>
<td>563.40 (SD = 180.85)</td>
</tr>
<tr>
<td>(4) Perfect-switch at participle</td>
<td>469.42 (SD = 152.47)</td>
<td>652.03 (SD = 190.00)</td>
<td>928.64 (SD = 372.60)</td>
</tr>
</tbody>
</table>

The same analyses conducted for the data obtained with the comprehension task were also conducted for the data gathered from the acceptability judgment task. Thus, these data were submitted to a 2 x 2 ANOVA with auxiliary type (progressive versus perfect) and switch site (switch at the auxiliary versus switch at the participle) as within-subjects factors. Separate analyses were conducted for each group of participants.

**Gaze duration.** For gaze duration, the early bilinguals’ results did not show a main effect of auxiliary type, $F(1,17) = .75, p = .398$, but they did show a main effect of switch site, $F(1,17) = 21.58, p < .001$. No by-participant interaction of auxiliary type and switch site was found, $F(1,17) = 2.87, p = .108$. Thus, for this reading measure, early bilinguals read switches at the
auxiliary more quickly than switches at the participle, but they did not display any reading time differences with respect to the particular auxiliary used in the sentence.

**Regression path time.** The regression path time results presented a main effect of auxiliary type, $F(1,17) = 7.26, p = .015$, and a main effect of switch site, $F(1,17) = 15.15, p = .001$, but no by-participant interaction of auxiliary type and switch site, $F(1,17) = 1.55, p = .231$. Consequently, participles accompanied by the progressive auxiliary were read more quickly than those in sentences that contained the perfect auxiliary. Participles were also read more quickly in sentences that included a switch at the auxiliary, compared to sentences that included a switch at the participle. However, there were no significant reading time differences between the two codeswitched conditions with the progressive structure (Conditions 1 and 2) nor between the two codeswitched conditions with the perfect structure (Conditions 3 and 4).

**Total time.** The results for total time are similar to those for regression path time. There was a main effect of auxiliary type, $F(1,17) = 4.56, p = .048$, and a main effect of switch site, $F(1,17) = 29.67, p = .048$. However, there was no by-participant interaction of auxiliary type and switch site, $F(1,17) = 2.87, p = .108$. Therefore, the early bilinguals exhibited significant total reading time differences of the participle between sentences that included the progressive structure and those that included the perfect structure. They also read participles significantly more quickly when they appeared in sentences with a switch at the auxiliary than when they appeared in sentences with a switch at the participle. However, no significant reading time differences arose when comparing the two crucial pairs of conditions.
3.4.2.2 Late bilinguals

The mean gaze duration, regression path time, and total time by condition for the group of late bilinguals is displayed in Table 3-8, with the standard deviations for the means shown in parentheses.

Table 3-8: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the late bilinguals during the acceptability judgment task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-</td>
<td>417.61</td>
<td>615.33</td>
<td>955.88</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>(SD = 170.33)</td>
<td>(SD = 304.41)</td>
<td>(SD = 523.71)</td>
</tr>
<tr>
<td>(2) Progressive-</td>
<td>425.80</td>
<td>586.91</td>
<td>940.40</td>
</tr>
<tr>
<td>Switch at participle</td>
<td>(SD = 110.84)</td>
<td>(SD = 238.32)</td>
<td>(SD = 405.07)</td>
</tr>
<tr>
<td>(3) Perfect-</td>
<td>421.06</td>
<td>680.51</td>
<td>1015.28</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>(SD = 127.53)</td>
<td>(SD = 373.50)</td>
<td>(SD = 639.87)</td>
</tr>
<tr>
<td>(4) Perfect-</td>
<td>475.10</td>
<td>920.44</td>
<td>1424.81</td>
</tr>
<tr>
<td>Switch at participle</td>
<td>(SD = 180.57)</td>
<td>(SD = 559.80)</td>
<td>(SD = 646.03)</td>
</tr>
</tbody>
</table>

**Gaze duration.** For gaze duration, the late bilinguals' results displayed no main effect of auxiliary type, $F(1,17) = 1.72, p = .208$, no main effect of switch site, $F(1,17) = 1.82, p = .195$, and no by-participant interaction of auxiliary type and switch site, $F(1,17) = .85, p = .370$. These results show that during this early reading measure there were no significant participle reading time differences among any of the four experimental conditions.

**Regression path time.** For regression path time, there was only a main effect of auxiliary type, $F(1,17) = 11.34, p = .004$. There was no main effect of switch site, $F(1,17) = 3.86, p = .066$, nor was there a by-participant interaction of auxiliary type and switch site, $F(1,17) = 3.72, p = .071$. Thus, overall, late bilinguals read participles more quickly when they were part of a progressive structure than when they were part of a perfect structure. However, there were no significant differences between conditions with respect to the particular point at which a switch occurred.
**Total time.** For total time, the results indicated a main effect of auxiliary type, $F(1,17) = 5.41$, $p = .033$, a main effect of switch site, $F(1,17) = 10.14$, $p = .005$, and a by-participant interaction of auxiliary type and switch site, $F(1,17) = 7.44$, $p = .014$. Follow-up paired-samples $t$ tests indicated significant mean differences between Conditions 3 and 4, $t(17) = 3.65$, $p = .002$, but not between Conditions 1 and 2, $t(17) = .18$, $p = .857$. In other words, the late bilinguals’ total reading time of the past participles was significantly slower when the switch occurred at the participle than when it occurred at the auxiliary, but no significant differences of total reading time of the present participles were found between Conditions 1 and 2.

### 3.5 Discussion

The results will be interpreted with respect to each of the research questions stated in section 3.1. The first question asked whether more frequently produced types of codeswitches (i.e., *estar*+English participle switches) were easier to process during reading comprehension than less frequent types of codeswitches (i.e., *haber*+English participle switches). This question can be addressed by comparing the mean differences between the two conditions in which the switch occurs at the participle with their respective baseline conditions, in which the switch occurs at the auxiliary. In other words, on the one hand, the means of Conditions 1 (Progressive-Switch at auxiliary) and 2 (Progressive-Switch at participle) are compared, and, on the other, the means of Conditions 3 (Perfect-Switch at auxiliary) and 4 (Perfect-Switch at participle) are compared. These differences are revealed by the interactions mentioned in section 3.4, which are summarized along with the other results in Table 3-9.
In the comprehension task, the fixation durations on the participle of both groups of participants were significantly longer for Condition 4 than for Condition 3. However, there were no significant differences between the fixation durations on the participle for Condition 1 and Condition 2. This was the case in the measures of early processing (i.e., gaze duration and regression path time) and the measure of later processing (i.e., total time). The fact that the early bilinguals exhibited a significant interaction with the measure of gaze duration and a marginally significant interaction with the measure of regression path time whereas the late bilinguals displayed a significant interaction only in regression path time simply demonstrates that both groups of participants may use different processing strategies when they encounter difficulty during reading. The late bilinguals seem to produce shorter initial fixations and then re-read when processing is costly while the early bilinguals seem to pause and produce longer fixations.

Table 3-9: Summary of the State College eye-tracking results by reading measure, participant group, and task

<table>
<thead>
<tr>
<th>Comprehension task</th>
<th>Early bilinguals</th>
<th>Late bilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main effect of auxiliary type</td>
<td>Main effect of switch site</td>
</tr>
<tr>
<td>Gaze</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Regression</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Total</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acceptability judgment task</th>
<th>Early bilinguals</th>
<th>Late bilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main effect of auxiliary type</td>
<td>Main effect of switch site</td>
</tr>
<tr>
<td>Gaze</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Total</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

√ Indicates a statistically significant effect ($p \leq .05$).
√* Indicates a marginally statistically significant effect ($p = .053$).
as soon as difficulty is encountered before re-reading. Although there were slight differences regarding the particular reading measure in which the interaction was found, both groups experienced processing difficulties with the haber+English participle switches from early stages of processing. Therefore, during reading comprehension, it was easier for participants to process the more frequent estar+English participle switches than the less common haber+English participle switches, when each of these were compared to their counterparts in which the switch occurred at the auxiliary.

The second research question concerned the role that age of L2 acquisition played during the comprehension of these types of codeswitches. This question can be addressed by examining the differences in processing costs between both groups of participants. A general result that arose is that the late bilinguals’ fixation durations on the participle tended to be longer than those of the early bilinguals’ overall (compare the means in Tables 3-4 and 3-7 for the early bilinguals with those in Tables 3-5 and 3-8 for the late bilinguals). Therefore, it seems that, generally speaking, these codeswitched sentences were slightly more costly for the late bilinguals to process than they were for the early bilinguals. There were also some differences between the two groups of participants with respect to the particular effects displayed in Table 3-9.

As mentioned above, in the comprehension task, both groups of participants showed very similar processing costs regarding the interaction of auxiliary type and switch site. Both groups of participants displayed the interaction in reading measures of early processing and in the reading measure of later processing. However, in the acceptability judgment task, differences arose between the two groups of participants. The early bilinguals presented main effects of auxiliary type and switch site during measures of early and later processing, showing that they had an easier time processing sentences with progressive structures and sentences with switches at the auxiliary phrase, but no significant interaction of auxiliary type and switch site emerged. The late bilinguals exhibited only a main effect of auxiliary type during the measures of early
processing, demonstrating that sentences with progressive structures were processed more easily than those with perfect structures. However, in the measure of later processing, a significant interaction of auxiliary type and switch site appeared. Therefore, when performing an acceptability judgment task, the early bilinguals always had more difficulty processing sentences with switches at the participle compared to those with switches at the auxiliary, regardless of the type of auxiliary involved (estar ‘be’ or haber ‘have’). On the other hand, although their initial processing patterns displayed very reduced effects of any kind, in the measure of total time, the late bilinguals only displayed processing costs with haber+English participle switches, but not withestar+English participle switches. To recapitulate, the two groups of participants displayed processing similarities of these types of codeswitches during the comprehension task, but the acceptability judgment task brought about differences in the way the early and late bilinguals read these codeswitches.

The third research question addressed the issue of whether, and if so, how different tasks influenced the way participants processed the different types of codeswitches. This question is answered by examining the differences in processing costs that each group of participants exhibits between the two tasks. The group of early bilinguals experienced a very visible effect of task, as displayed by the presence of a significant interaction of auxiliary type and switch site in all measures of processing in the comprehension task contrasted with the absence of an interaction in the acceptability judgment task. In the comprehension task, the early bilinguals were very sensitive to the differences betweenestar+English participle switches and haber+English participle switches. During early and later measures of processing, they processed the former type of switches with more ease than the latter type, when each of these was compared to its corresponding condition in which the switch occurred at the auxiliary. However, in the acceptability judgment task, this effect disappeared completely. That is, the early bilinguals displayed more processing costs with sentences that included a switch at the participle, regardless
of the auxiliary involved. Therefore, with this group of participants, only the processing costs that they displayed during the comprehension task reflect what is found in natural production. Thus, for the early bilinguals, when the task was to simply try to comprehend what they were reading, their comprehension difficulty mirrored the production frequency of these two types of switches. The *estar*+English participle switches, to which they were probably more exposed in natural production, were more easily processed by them than the *haber*+English participle switches, to which they were probably almost never exposed. Nonetheless, when the early bilinguals were given the task of judging the codeswitched sentences, they seemed to engage in metalinguistic mechanisms that differed from more natural processing mechanisms, leading them to process *estar*+English participle switches and *haber*+English participle switches with similar difficulty.

With the group of late bilinguals, however, the story is different. In this case, the interaction of auxiliary type and switch site was present in both the comprehension task and the acceptability judgment task. The interaction arose in an earlier reading measure (i.e., regression path time) in the comprehension task, but it was present during the measure of total time, regardless of the task given to the participants. In other words, in the acceptability judgment task, it took the late bilinguals a little longer to display the different processing of *estar*+English participle switches and *haber*+English participle switches, but it surfaced in both tasks, nevertheless. This result suggests that task effects did not influence the processing of all participants in the same way.

How can these differences between the two groups of participants be clarified? A possible explanation may have to do with awareness of the social stigma that is often attached to the linguistic act of codeswitching. On the one hand, the early bilinguals, who have been exposed to codeswitching from an early age, may have been brought up with the idea that, although bilinguals engage in codeswitching, it is something that should be avoided because it is ultimately
“wrong” to mix languages. Therefore, when they are asked to read the codeswitched sentences for comprehension, their processing costs mirror natural production. However, when they are asked to read in order to judge the acceptability of the sentences, their awareness of the stigma attached to codeswitching may lead them to process the codeswitches differently, by focusing on them in a way that they normally would not and, as a result, by displaying processing difficulties for any type of complex codeswitches, such as those that take place within the auxiliary phrase, regardless of the fact that they probably use and are more exposed to some of those switches (i.e., \textit{estar}+English participle switches) more than others (\textit{haber}+English participle switches). On the other hand, the late bilinguals, who acquired their L2 at a later age and whose exposure to codeswitching is therefore reduced, may not be completely aware of the stigma that is generally attached to codeswitching. Therefore, when they are asked to read the sentences in order to judge the acceptability of the codeswitches, their processing is probably more neutral and unbiased, that is, less informed by the social rejection of codeswitching and simply influenced by what they have been exposed to from the time they acquired their L2 and began codeswitching with other bilinguals. This may be why the late bilinguals processed the codeswitches similarly in both the comprehension task and the acceptability judgment task.

This reasoning is only speculative and requires closer examination. If this were in fact the case, one would expect the participants’ overt judgments to reflect their processing costs. However, this is not what always occurred. As mentioned above, the late bilinguals displayed an interaction of auxiliary type and switch site in the reading measure of total time, therefore showing that \textit{estar}+English participle switches were more easily processed by them than \textit{haber}+English participle switches. Their judgments also displayed the preference of \textit{estar}+English participle switches over \textit{haber}+English participle switches. Overall, the sentences belonging to Conditions 1, 2, and 3 were judged acceptable over 90% of the time. The sentences belonging to Condition 4 were only judged acceptable about 64% of the time. Therefore, the
processing costs of the late bilinguals coincided with their acceptability judgments. Based on the early bilinguals’ lack of interaction of auxiliary type and switch site as they processed the participles, it is reasonable to also expect these results to concur with their judgments. Thus, one would expect all sentences containing switches at the participle (i.e., those belonging to Conditions 2 and 4) to be judged unacceptable by the early bilinguals, regardless of the auxiliary involved. However, this was not the case. As with the late bilinguals, with the early bilinguals, sentences belonging to Conditions 1, 2, and 3 were judged acceptable over 90% of the time, but sentences belonging to Condition 4 were judged acceptable only about 80% of the time. Although there is less of a difference between Condition 3 and Condition 4 regarding the proportion of acceptable judgments provided by the early bilinguals, the sentences belonging to Condition 4 still received the lowest proportion of acceptable judgments.

Regarding the comprehension task, one could expect the processing difficulties to also be reflected in the participants’ overt responses to the comprehension questions. Because both participant groups displayed the longest reading times and, thus, the largest processing difficulties with sentences belonging to Condition 4, it is reasonable to expect the questions belonging to this condition to impose the most problems for participants. Nonetheless, for both groups of participants, the percentage of correct responses to questions belonging to Condition 4 were no different from the percentage of correct responses to questions belonging to any of the other three conditions. Therefore, although the participants exhibited processing difficulties with haber+English participle switches, which were captured by the very precise eye-tracking reading measures, these difficulties did not impede their overall comprehension of the codeswitched sentences.
3.6 Conclusion

This study examined whether more frequently produced codeswitches were easier to process than less frequently produced codeswitches. The study was also aimed at examining the influence of age of L2 acquisition and of task on the way that Spanish-English bilinguals processed frequent and infrequent types of codeswitches. The results showed that frequent codeswitches were processed with significantly greater ease than less frequent codeswitches by both groups of participants. These findings lend support for constraint satisfaction models of language processing, such as the Production-Distribution-Comprehension approach (Gennari & MacDonald, 2009; MacDonald & Thornton, 2009), in which linguistic experience plays a crucial role in the way language is processed. According to these models, the more exposure individuals have to particular linguistic structures, the easier they are for them to process the next time they encounter them. In this case, both groups of Spanish-English bilinguals are probably more exposed to the more frequent estar+English participle switches than they are to the less common haber+English participle switches, and this difference in exposure is reflected in their processing costs while reading these two types of switches. The results described above were found when participants were asked to read for comprehension. However, when participants were asked to read in order to provide judgments, the results were different, particularly in the case of the early bilinguals. With this task, the processing costs of the early bilinguals no longer reflected natural production patterns. The late bilinguals’ processing costs still reflected natural production patterns, but only in the measure of later processing. Therefore, at least to a certain degree, the task given to participants affected the way they processed the codeswitched sentences. The three questions examined in this study were further addressed with participants who belong to an established community of codeswitchers. The methods and the findings are described in Chapter 4.
Chapter 4

Comprehension costs associated with auxiliary phrase switches in an established community of codeswitchers

The linguistic act of codeswitching is strongly linked to the speech community, which can be described as a regionally or socially definable group characterized by the use of a shared linguistic system and shared sociolinguistic norms (Labov, 1972). One of the main reasons why bilinguals engage in codeswitching is precisely to emphasize group membership in bilingual communities. Codeswitching with this purpose is particularly common in immigrant minority groups and is used to express solidarity. In this sense, codeswitching becomes a marker of bilingual social identity and indicates dual affiliation with two cultures (Heller, 1988, 1995). Most of the reports that have provided descriptions of the codeswitching patterns that Spanish-English bilinguals use in spontaneous speech come from established communities of bilingual codeswitchers (Deuchar et al., 2012; Lipski, 1985; Pfaff, 1979; Poplack, 1980).

The two groups of bilinguals who completed the study described in Chapter 3 were comprised of students at The Pennsylvania State University, located in State College, Pennsylvania. Although many of them may come from Hispanic countries or from other cities in the United States (U.S.) with larger Hispanic communities, at the moment they participated in the experiment, they had been living in the State College area for some time. State College is a college town in the center of Pennsylvania, where the Hispanic and Latino population nearly reaches 4% (U.S. Census Bureau, 2010c). The student body at the Penn State University Park campus is largely comprised by White students (71%), while the Hispanic students make up 4.7% of the student population (Penn State University Budget Office, 2011). The Spanish-English bilinguals who completed the State College study are definitely more isolated than other
bilinguals in larger U.S. communities where codeswitching is more of a common practice. They are probably not as exposed to codeswitching and, by extension, probably do not codeswitch as much as bilinguals who belong to a larger community of codeswitchers. It is also possible that the codeswitching patterns followed by the Spanish-English bilinguals in State College differ from those in a larger codeswitching community. To this respect, codeswitching studies (e.g., Poplack, 1985) have shown that codeswitching norms or strategies can differ from one bilingual community to another.

It is important, therefore, to examine whether the correspondence between production patterns and comprehension costs that was found in the State College study also obtains in a more established codeswitching community. With this idea in mind, the same eye-tracking experiment that was described in Chapter 3 was also conducted with Spanish-English bilinguals in an established community of bilingual codeswitchers. Poplack (1980) collected her data from Spanish-English bilinguals in East Harlem, an area also known as Spanish Harlem or El Barrio. The balanced bilinguals in her study tended to produce intrasentential switches, which are the type of switches used in this eye-tracking study. Based on Poplack’s findings, I collected data from students attending The City College of New York (CCNY), which is located in West Harlem. The racial composition of this city and university differs drastically from that of State College and Penn State University. According to the 2010 census, the Hispanic and Latino population of New York City reaches 28.6% (U.S. Census Bureau, 2010b). Moreover, in CCNY, Hispanics constitute the largest ethnic group, representing approximately 32% of the student body whereas White students make up about 24% (The City College of The City University of New York, n.d.). By being part of such a large community of Spanish-English bilingual codeswitchers, it is possible that these bilingual participants are more exposed to the codeswitching patterns examined in this study.
4.1 Research questions

The research questions addressed in this study were the same as the ones addressed in the previous study. They are re-stated below.

1) Are more frequently produced types of codeswitches easier to process by the comprehension system than less frequent types of codeswitches?

2) Does age of second language (L2) acquisition play a role in the comprehension of these types of switches?

3) Do different tasks influence processing patterns during the comprehension of these codeswitches?

4.2 Method

4.2.1 Participants

Approximately half of the data for this study was collected between March and May 2011 and the rest of the data was collected between September and November 2011. Sixty-nine participants contributed to this study. However, ten participants were excluded from the study for different reasons. Two participants were excluded because of technical problems with the eye-tracking equipment and four additional participants were eliminated because they reported that they hardly ever codeswitched. The last four participants were excluded because their mean gaze durations on the critical region (i.e., the participle) in one or more of the four experimental conditions did not reach 100 milliseconds (ms) or exceeded 1500 ms. The remaining 59 participants were divided into two groups: a group of 42 early bilinguals and a group of 17 late bilinguals. Participants were placed into one of these two groups based on the characteristics and scores obtained from several measures of language proficiency mentioned below.
All participants were asked to complete the same measures of language proficiency that were used in the previous study: the Language History Questionnaire (Appendix A), the Boston Naming Vocabulary Test in English and Spanish (Appendix B), and the two grammatical competency tests (MELICET for English and DELE for Spanish, Appendix C). All participants were undergraduate or graduate students at CCNY, with a very wide range of majors of study. Approximately 78% of the participants lived in Harlem, in Washington Heights, in the Bronx, or in Queens, which constitute areas of New York City with a sizeable Hispanic population. The remaining 22% lived in other boroughs of New York City or in other suburbs of New York. Moreover, 63% of the participants were Dominican (either born in the Dominican Republic or raised in a Dominican family in New York). The other reported Hispanic nationalities were, in decreasing order, Ecuadorian, Colombian, Mexican, Honduran, Nicaraguan, Spanish, Puerto Rican, and Guatemalan. Some of the participants reported a mixed origin among these Hispanic nationalities.

For all participants the native language of their parents and siblings was Spanish. When asked about the languages that were spoken in their home, approximately half of the participants in each group reported that both English and Spanish were spoken while the other half said that only Spanish was spoken in their home. Regarding the acquisition of each language, on average, the early bilinguals acquired both languages at similar ages, while the late bilinguals acquired their second language (English) at a later age. In this respect, half of the participants in the group of early bilinguals reported speaking both Spanish and English before reaching five years of age. However, none of the participants in the group of late bilinguals began speaking English before age 5.

The early bilinguals had been living in the U.S. for nine years or more at the time of the experiment while the late bilinguals had been living in the U.S. for eight years or less. Both participant groups reported using and being exposed to English and Spanish on a regular basis.
Overall, participants reported being more exposed to English than Spanish while reading and while watching television or movies. In general, the group of early bilinguals reported being equally exposed to Spanish and English while listening to music and the radio whereas the group of late bilinguals reported being more exposed to Spanish music and radio. Regarding exposure to both languages with family and friends, the tendency in both groups was that with family members they were either more exposed to Spanish or exposed equally to both languages. With friends, they were either more exposed to English or exposed equally to both languages.

Participants were also asked about their knowledge of languages other than English and Spanish. Approximately 30% of participants had taken or were taking language classes of Italian, French, Portuguese, or Japanese. In addition, one participant said that she had learned some Quichua from her grandfather.

As part of the language questionnaire, participants also answered several questions about their codeswitching practices. They all reported engaging in oral codeswitching occasionally, frequently, or constantly. When asked with whom they tended to codeswitch, participants’ most frequent responses were with family members, with friends, or with other fluent Spanish-English bilinguals. Moreover, nearly 85% of the participants reported engaging in written codeswitching in environments such as email, online instant messages, online chat rooms, and text messages.

Participants were also inquired about the directionality of their codeswitching acts. Approximately one third of the participants reported switching more often from Spanish to English; another third of the participants said that they tended to switch from English to Spanish; an additional third reported equally switching in both directions. A small minority indicated being unsure of their preferred switching directionality. Finally, when asked about their reasons for codeswitching, participants’ most frequent responses related to quicker access to words or phrases in one language over the other, to better conveyance of meaning when using words or phrases in a particular language, and to ease, comfort, or habit. Concerning the language
proficiency measures, the group of early bilinguals provided higher self-ratings of English proficiency and received higher scores on the English vocabulary and grammar tests. The group of late bilinguals had higher self-ratings and test scores for Spanish proficiency. Independent-samples $t$ tests were conducted to determine the presence of significant differences between both groups of participants for these proficiency measures. The results displayed significant differences between the early bilinguals and the late bilinguals for almost all measures (self-ratings for English proficiency: $t(57) = 2.19, p = .033$; self-ratings for Spanish proficiency: $t(57) = 4.06, p < .001$; BNT score for English: $t(57) = 2.74, p = .008$; BNT score for Spanish: $t(57) = 6.48, p < .001$; DELE score: $t(57) = 5.21, p < .001$). The exception was the MELICET score, for which no significant difference between groups was found, $t(57) = .19, p = .850$. Table 4-1 displays the means and ranges of different characteristics by participant group.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Early bilinguals n = 42</th>
<th>Late bilinguals n = 17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>21 (18-33)</td>
<td>22 (18-32)</td>
</tr>
<tr>
<td>Age of arrival in U.S.</td>
<td>4 (0-20)</td>
<td>17 (10-32)</td>
</tr>
<tr>
<td>Years spent living in U.S.</td>
<td>17 (9-22)</td>
<td>5 (0.5-8)</td>
</tr>
<tr>
<td>Age of reported achieved fluency in English</td>
<td>9 (2-22)</td>
<td>15 (8-21)</td>
</tr>
<tr>
<td>Age of reported achieved fluency in Spanish</td>
<td>6 (2-16)</td>
<td>5 (2-7)</td>
</tr>
<tr>
<td>Difference in years between reported achieved fluency in L1 and L2</td>
<td>5 (0-17)</td>
<td>11 (1-16)</td>
</tr>
<tr>
<td>Self-ratings for English proficiency (/10)</td>
<td>9 (5.75-10)</td>
<td>8.4 (6.5-10)</td>
</tr>
<tr>
<td>Self-ratings for Spanish proficiency (/10)</td>
<td>8.3 (5.75-10)</td>
<td>9.5 (8.25-10)</td>
</tr>
<tr>
<td>BNT score for English (/30)</td>
<td>20 (9-27)</td>
<td>16 (9-22)</td>
</tr>
<tr>
<td>BNT score for Spanish (/30)</td>
<td>14 (7-24)</td>
<td>21 (17-26)</td>
</tr>
<tr>
<td>MELICET score (/50)</td>
<td>40 (30-50)</td>
<td>40 (30-46)</td>
</tr>
<tr>
<td>DELE score (/50)</td>
<td>35 (20-45)</td>
<td>43 (34-48)</td>
</tr>
</tbody>
</table>
4.2.2 Materials and Design

The codeswitched experimental stimuli used in the study were the same as those used in the State College study (Appendix D). In this study, however, a series of English and Spanish sentences were added to the experiment and presented to the participants in separate language blocks. The English and Spanish blocks served the purpose of providing a general measure of reading speed for each participant in both languages. The results of this measure will be discussed in the section 4.4.3. Including these two additional language blocks also ensured that participants were equally exposed to both languages throughout the experiment. The experimental sentences in the English block had the same structure as the experimental codeswitched sentences. In other words, they were identical to the codeswitched sentences, with the exception that the beginning of the codeswitched sentences was translated to English for the unilingual English sentences. In Table 4-2, two examples of the unilingual English sentences are displayed alongside the codeswitched sentences from Table 3-2 for comparison. All unilingual English sentences are included in Appendix E.

Table 4-2: Examples of unilingual English sentences (matched translations of the codeswitched sentences)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Sample sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressive-Unilingual</td>
<td>The director confirmed that the actors are <em>rehearsing</em> their lines for the movie.</td>
</tr>
<tr>
<td>(1) Progressive-Switch at auxiliary</td>
<td><em>El director confirmó que los actores</em> are <em>rehearsing</em> their lines for the movie.</td>
</tr>
<tr>
<td>(2) Progressive-Switch at participle</td>
<td><em>El director confirmó que los actores están rehearsing</em> their lines for the movie.</td>
</tr>
<tr>
<td>Perfect-Unilingual</td>
<td>The director confirmed that the actors have <em>rehearsed</em> their lines for the movie.</td>
</tr>
<tr>
<td>(3) Perfect-Switch at auxiliary</td>
<td><em>El director confirmó que los actores</em> have <em>rehearsed</em> their lines for the movie.</td>
</tr>
<tr>
<td>(4) Perfect-Switch at participle</td>
<td><em>El director confirmó que los actores han rehearsed</em> their lines for the movie.</td>
</tr>
</tbody>
</table>
This particular design allowed for an additional comparison between the English sentences and the codeswitched sentences that could reveal the presence of switch costs, if any. The results of this additional comparison will be discussed in Chapter 5.

The experimental sentences in the Spanish block also had the same structure as the experimental sentences in the English and Codeswitched blocks. That is, the unilingual Spanish sentences included subordinate clauses with progressive and perfect structures, but they were not direct translations of the English experimental sentences. Two examples, one with the progressive structure and one with the perfect structure, are displayed below.

(35) *El supervisor dijo que los técnicos están arreglando las impresoras de la compañía.*

‘The supervisor said that the technicians are fixing the printers of the company.’

(36) *La reportera anunció que los senadores han aprobado las leyes de conservación ambiental.*

‘The reporter announced that the senators have approved the laws of environmental conservation.’

Like the codeswitched sentences, the unilingual English sentences were divided into six reading files, each of which included 16 experimental unilingual sentences (eight sentences per condition). Each reading file also included 16 filler sentences, which had the same overall length as the experimental items. Two examples of the fillers are provided below.

(37) During the meeting, Henry was giving a presentation while Barbara was taking notes.

(38) Chris and Matthew had been playing cards on the street corner all day long.

The same care that was taken with respect to task demands in the Codeswitched block was also taken in the English block. For the comprehension block, the questions for half of the sentences required a “yes” response while those for the other half required a “no” response. Furthermore, half of the questions obliged participants to focus on the beginning of the sentence whereas the other half required attention to the end of the sentence. For the acceptability judgment task, half of the filler sentences included a grammatical error towards the beginning of the sentence and the
other half included the grammatical error towards the end of the sentence. Two examples of the ungrammatical fillers are exhibited below.

39) *Sean will be going to these mall to buy dinner and watch the new movie.

(40) *Scott would have finished his homework if he didn’t have to goes to work.

In the case of the Spanish block, only one reading file for each task was created. Therefore, all participants read the same sentences in the Spanish comprehension block and in the Spanish acceptability judgment block. Each of these blocks included 16 experimental sentences, such as those displayed in examples (35) and (36). They also included 16 filler sentences, which were similar to the English fillers with respect to overall length and linguistic structure. The same control measures for the materials described for the codeswitched sentences and the English sentences were also followed for the Spanish sentences and their respective questions in each task block. Two examples of the Spanish fillers are displayed below. Example (41) belongs to the Spanish comprehension block and example (42) belongs to the Spanish acceptability judgment block.

(41) Dentro de cuatro meses Miguel ya se habrá recuperado totalmente de su operación.

`In four months Michael will already have recuperated totally from his operation.'

(42) *Marcos estará participando en una carrera que se llevará a cabo el semana que viene.

`Marcus will be participating in a race that will take place the next week.'

The sentences within each of the English and Spanish reading files were pseudo-randomized to avoid order effects. Language block presentation order and task block presentation order within each language block were counterbalanced across participants.
4.2.3 Procedure

The procedure used in this study was identical to that used in the State College study. However, the addition of the English and the Spanish language blocks increased the length of the experiment considerably. Therefore, in order to avoid participant fatigue, the experiment was divided in two 1.5-hour sessions that participants completed on two separate days. Half of the reading blocks and half of the proficiency measures were completed in one session and the remaining parts of the experiment were completed in the second session. Participants were compensated $15 per session for their participation.

4.3 Predictions

The predictions in this study were the same as those described in the previous study. To briefly restate them here, regarding the first research question, I expected the more frequent *estar*+English participle switches to be processed more easily than the less frequent *haber*+English participle switches. In other words, when each of these conditions was compared to its respective baseline condition in which the switch occurred at the auxiliary, I expected to find different results depending on the particular linguistic structure involved. Mainly, I expected to find no significant reading time differences between the two conditions with the progressive structure. However, I expected the reading times for *haber*+English participle switches to be significantly larger than those for the perfect structure condition in which the switch occurred at the auxiliary. Because these bilinguals were part of a more established codeswitching community, I expected them to be even more sensitive to the asymmetrical frequency in natural production of the *estar*+English participle switches and the *haber*+English participle switches. With respect to the second research question, I expected to find processing differences between
the early and the late bilinguals based on the differences between them in terms of age of L2 acquisition and exposure to codeswitching patterns. Finally, regarding the third research question, I expected reading times on the critical region to be significantly affected by the particular task that participants were asked to complete.

4.4 Results

Data were extracted and analyzed in the same way as they were in the State College study. The two reading measures of early processing (i.e., gaze duration and regression path time) and the measure of later processing (i.e., total time) were extracted for the participle in the experimental sentences. Analyses were conducted separately for each task and for each participant group. First, the results pertaining to the comprehension task will be discussed, followed by those for the acceptability judgment task. The results pertaining to the participants’ reading speed in both languages will be presented at the end of this section.

4.4.1 Results for the comprehension task

Table 4-3 presents the proportions of correct responses to the comprehension questions by condition and participant group.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent of correct answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early bilinguals</td>
<td>Late bilinguals</td>
</tr>
<tr>
<td>(1) Progressive-Switch at auxiliary</td>
<td>89.58%</td>
</tr>
<tr>
<td>(2) Progressive-Switch at participle</td>
<td>91.07%</td>
</tr>
<tr>
<td>(3) Perfect-Switch at auxiliary</td>
<td>90.77%</td>
</tr>
<tr>
<td>(4) Perfect-Switch at participle</td>
<td>92.26%</td>
</tr>
</tbody>
</table>
The proportions of correct responses are very similar between both groups of participants and across all four conditions. Overall, both groups of participants answered most of the comprehension questions correctly. Therefore, the results show that participants were paying attention and that they were able to understand the sentences that they read.

### 4.4.1.1 Early bilinguals

Table 4-4 presents the mean gaze duration, regression path time, and total time by condition for the group of early bilinguals. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration (SD)</th>
<th>Regression path time (SD)</th>
<th>Total time (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-Switch at auxiliary</td>
<td>347.23 (136.57)</td>
<td>525.76 (273.69)</td>
<td>691.38 (314.60)</td>
</tr>
<tr>
<td>(2) Progressive-Switch at participle</td>
<td>384.41 (128.63)</td>
<td>606.11 (299.96)</td>
<td>750.15 (265.51)</td>
</tr>
<tr>
<td>(3) Perfect-Switch at auxiliary</td>
<td>316.06 (101.32)</td>
<td>517.82 (272.32)</td>
<td>649.57 (247.68)</td>
</tr>
<tr>
<td>(4) Perfect-Switch at participle</td>
<td>427.18 (193.00)</td>
<td>698.91 (321.49)</td>
<td>944.10 (442.78)</td>
</tr>
</tbody>
</table>

A two-way repeated measures analysis of variance (ANOVA) was performed in order to evaluate the effect of auxiliary type and switch site on gaze duration, regression path time, and total time. Auxiliary type (progressive versus perfect) and switch site (switch at the auxiliary versus switch at the participle) were the within-subjects factors.

**Gaze duration.** For gaze duration, the results yielded a main effect of switch site, $F(1,41) = 18.04, p < .001$, but not of auxiliary type, $F(1,41) = .26, p = .616$. There was also a significant by-participant interaction of auxiliary type and switch site, $F(1,41) = 5.30, p = .026$. Paired-samples $t$ tests were conducted to follow up the significant interaction. They indicated significant mean differences between Conditions 3 and 4, $t(41) = 4.67, p < .001$. Nonetheless, no significant
mean differences were found between Conditions 1 and 2, \( t(41) = 1.57, p = .123 \). Therefore, during the measure of gaze duration, the early bilinguals read \( haber+English \) participle switches significantly more slowly than perfect structures in which the switch occurred at the auxiliary. However, \( estar+English \) participle switches and their corresponding phrasal boundary switches were read at a similar speed.

**Regression path time.** For regression path time, the results only displayed a main effect of switch site, \( F(1,41) = 17.05, p < .001 \). No main effect of auxiliary type, \( F(1,41) = 2.59, p = .115 \), and no by-participant interaction of auxiliary type and switch site, \( F(1,41) = 2.03, p = .162 \), were found. During this reading measure, all sentences in which the switch occurred at the auxiliary were read more quickly than sentences in which the switch occurred at the participle, regardless of the particular auxiliary used.

**Total time.** For total time, the results showed a main effect of auxiliary type, \( F(1,41) = 7.41, p = .009 \), a main effect of switch site, \( F(1,41) = 27.87, p < .001 \), and a significant by-participant interaction of auxiliary type and switch site, \( F(1,41) = 9.49, p = .004 \). Subsequent pairwise contrasts indicated significant mean differences between Conditions 3 and 4, \( t(41) = 5.22, p < .001 \), but not between Conditions 1 and 2, \( t(41) = 1.32, p = .194 \). During this measure of later processing, the early bilinguals read the two codeswitched conditions with the progressive structure at a similar speed. On the contrary, switches at the past participle were read significantly more slowly than switches at the perfect auxiliary.

### 4.4.1.2 Late bilinguals

The mean gaze duration, regression path time, and total time by condition for the group of late bilinguals is displayed in Table 4-5. Standard deviations are displayed in parentheses.
For gaze duration, the results for the late bilinguals only indicated a main effect of switch site, $F(1,16) = 9.40, p = .007$. There was no main effect of auxiliary type, $F(1,16) = .76, p = .396$, nor was there a by-participant interaction of auxiliary type and switch site, $F(1,16) = .17, p = .690$. Therefore, during this measure of early processing, the late bilinguals read all participles in sentences with phrasal boundary switches more quickly than those with switches at the present or past participles.

**Regression path time.** The regression path time results exhibited a main effect of switch site, $F(1,16) = 11.49, p = .004$, but no main effect of auxiliary type, $F(1,16) = 3.26, p = .090$. Moreover, there was a significant by-participant interaction of auxiliary type and switch site, $F(1,16) = 6.69, p = .020$. Follow-up paired-samples $t$ tests exhibited significant mean differences between Conditions 3 and 4, $t(16) = 4.40, p < .001$, but not between Conditions 1 and 2, $t(16) = .79, p = .439$. The late bilinguals read perfect structures with phrasal boundary switches significantly more quickly than perfect structures with participle switches. However, they did not read the two types of switches involving the progressive structure at significantly different speeds.

**Total time.** The results of the ANOVA for total time were similar to those for regression path time. They lacked a main effect of auxiliary type, $F(1,16) = 2.46, p = .152$. Nonetheless, there was a main effect of switch site, $F(1,16) = 22.25, p < .001$, as well as a significant by-participant

Table 4-5: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the late bilinguals during the comprehension task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-</td>
<td>398.18 (SD = 166.97)</td>
<td>569.53 (SD = 187.72)</td>
<td>731.73 (SD = 259.37)</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Progressive-</td>
<td>457.90 (SD = 195.02)</td>
<td>623.06 (SD = 300.24)</td>
<td>808.67 (SD = 372.56)</td>
</tr>
<tr>
<td>Switch at participle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Perfect-</td>
<td>359.48 (SD = 124.77)</td>
<td>571.14 (SD = 272.85)</td>
<td>703.01 (SD = 250.94)</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Perfect-</td>
<td>449.21 (SD = 217.42)</td>
<td>849.17 (SD = 413.54)</td>
<td>1014.19 (SD = 361.38)</td>
</tr>
<tr>
<td>Switch at participle</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Gaze duration.** For gaze duration, the results for the late bilinguals only indicated a main effect of switch site, $F(1,16) = 9.40, p = .007$. There was no main effect of auxiliary type, $F(1,16) = .76, p = .396$, nor was there a by-participant interaction of auxiliary type and switch site, $F(1,16) = .17, p = .690$. Therefore, during this measure of early processing, the late bilinguals read all participles in sentences with phrasal boundary switches more quickly than those with switches at the present or past participles.
interaction of auxiliary type and switch site, $F(1,16) = 6.65, p = .020$. Subsequent pairwise contrasts only displayed significant mean differences between Conditions 3 and 4, $t(16) = 4.69, p < .001$. There were no significant differences between the means of Conditions 1 and 2, $t(16) = 1.38, p = .186$. For this reading measure, in sentences with the progressive structure, the late bilinguals read the participles at a similar speed, regardless of the location of the switch. In sentences with the perfect structure, however, participles were read at a slower speed when the switch occurred at the participle, compared to when it occurred at the auxiliary.

### 4.4.2 Results for the acceptability judgment task

Table 4-6 displays the proportions of acceptable judgments of the codeswitched sentences by condition and participant group.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent of acceptable judgments</th>
<th>Early bilinguals</th>
<th>Late bilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-Switch at auxiliary</td>
<td>89.29%</td>
<td>95.59%</td>
<td></td>
</tr>
<tr>
<td>(2) Progressive-Switch at participle</td>
<td>82.44%</td>
<td>91.18%</td>
<td></td>
</tr>
<tr>
<td>(3) Perfect-Switch at auxiliary</td>
<td>86.31%</td>
<td>93.38%</td>
<td></td>
</tr>
<tr>
<td>(4) Perfect-Switch at participle</td>
<td>63.69%</td>
<td>50.00%</td>
<td></td>
</tr>
</tbody>
</table>

Table 4-6 shows that, overall, Conditions 1, 2, and 3 received the highest proportions of acceptable judgments from both groups of participants. For these three conditions, the proportions of acceptable judgments from the early bilinguals were lower than those from the late bilinguals, but the differences between them do not reach ten percentage points. Both groups of participants judged the sentences that included *haber*+English participle switches as the least acceptable sentence type. The percentage of acceptable judgments for sentences belonging to
Condition 4 is particularly low in the case of the late bilinguals, who only considered them acceptable half of the time.

### 4.4.2.1 Early bilinguals

Table 4-7 presents the means and standard deviations for gaze duration, regression path time, and total time by condition for the group of early bilinguals.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Progressive-</td>
<td>338.93 (SD = 132.70)</td>
<td>581.45 (SD = 286.17)</td>
<td>778.12 (SD = 376.43)</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Progressive-</td>
<td>441.51 (SD = 144.75)</td>
<td>718.66 (SD = 355.45)</td>
<td>1130.14 (SD = 592.28)</td>
</tr>
<tr>
<td>Switch at participle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) Perfect-</td>
<td>330.86 (SD = 147.43)</td>
<td>589.72 (SD = 383.87)</td>
<td>819.83 (SD = 543.28)</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) Perfect-</td>
<td>444.06 (SD = 210.16)</td>
<td>859.82 (SD = 528.77)</td>
<td>1281.15 (SD = 783.31)</td>
</tr>
<tr>
<td>Switch at participle</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the acceptability judgment task, the data were also submitted to a 2 x 2 ANOVA with auxiliary type (progressive versus perfect) and switch site (switch at the auxiliary versus switch at the participle) as the within-subjects factors. Separate analyses were conducted for each group of participants.

**Gaze duration.** For gaze duration, the early bilinguals’ results only showed a main effect of switch site, $F(1,41) = 30.07, p < .001$. No main effect of auxiliary type, $F(1,41) = .03, p = .864$, and no by-participant interaction of auxiliary type and switch site was found, $F(1,41) = .12, p = .732$. Therefore, at this point in processing, all sentences that included a switch at the auxiliary were read significantly more quickly than sentences that included a switch at the participle, regardless of the particular structure involved.
Regression path time. The same results were found for regression path time. There was a main effect of switch site, $F(1,41) = 21.14, p < .001$, but no main effect of auxiliary type, $F(1,41) = 1.86, p = .180$, and no by-participant interaction of auxiliary type and switch site, $F(1,41) = 1.93, p = .172$. Once again, present and past participles were read more quickly by the early bilinguals when they appeared in sentences in which the switch occurred at the auxiliary than when they appeared in sentences in which the switch occurred at the participle.

Total time. The results for the two measures of early processing persisted in the measure of total time. Yet again the analyses yielded only a main effect of switch site, $F(1,41) = 52.74, p < .001$. However, they did not display a main effect of auxiliary type, $F(1,41) = 1.78, p = .190$, nor a by-participant interaction of auxiliary type and switch site, $F(1,41) = 1.66, p = .204$. Therefore, the early bilinguals always read the participles at a quicker speed when they appeared in sentences in which the switch occurred at the auxiliary, as compared to sentences in which the switch occurred at the participle.

4.4.2.2 Late bilinguals

The mean gaze duration, regression path time, and total time by condition for the group of late bilinguals is displayed in Table 4-8, with the standard deviations shown in parentheses.

Table 4-8: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the late bilinguals during the acceptability judgment task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-switch at auxiliary</td>
<td>434.68 (SD = 216.62)</td>
<td>773.77 (SD = 381.38)</td>
<td>959.43 (SD = 463.34)</td>
</tr>
<tr>
<td>(2) Progressive-switch at participle</td>
<td>484.04 (SD = 166.49)</td>
<td>725.32 (SD = 350.35)</td>
<td>993.10 (SD = 481.29)</td>
</tr>
<tr>
<td>(3) Perfect-switch at auxiliary</td>
<td>442.33 (SD = 136.81)</td>
<td>718.48 (SD = 351.73)</td>
<td>902.46 (SD = 332.90)</td>
</tr>
<tr>
<td>(4) Perfect-switch at participle</td>
<td>530.82 (SD = 242.89)</td>
<td>1244.65 (SD = 943.92)</td>
<td>1336.46 (SD = 763.35)</td>
</tr>
</tbody>
</table>
Gaze duration. For gaze duration, the late bilinguals’ results displayed no main effect of auxiliary type, $F(1,16) = .33, p = .571$, no main effect of switch site, $F(1,16) = 3.88, p = .066$, and no by-participant interaction of auxiliary type and switch site, $F(1,16) = .32, p = .578$. Therefore, during this reading measure of early processing there were no significant differences in participle reading times among any of the four experimental conditions.

Regression path time. For regression path time, there was a main effect of auxiliary type, $F(1,16) = 6.72, p = .020$, a main effect of switch site, $F(1,16) = 4.48, p = .050$, and a significant by-participant interaction of auxiliary type and switch site, $F(1,16) = 9.11, p = .008$. Subsequent pairwise contrasts displayed significant mean differences between Conditions 3 and 4, $t(16) = 2.76, p = .014$, but not between Conditions 1 and 2, $t(16) = .57, p = .575$. In other words, the late bilinguals’ reading times of the past participles were significantly longer when they were in sentences that included a switch at the participle, compared to when the participles were in sentences that included a switch at the auxiliary. However, no significant present participle reading time differences were found between the sentences that included a switch at the auxiliary and those that included a switch at the participle.

Total time. For total time, the ANOVA did not exhibit a main effect of auxiliary type, $F(1,16) = 2.61, p = .126$. It did, however, display a main effect of switch site, $F(1,16) = 8.75, p = .009$. There was also a significant by-participant interaction of auxiliary type and switch site $F(1,16) = 6.71, p = .020$. Follow-up paired-samples $t$ tests indicated significant mean differences between Conditions 3 and 4, $t(16) = 3.55, p = .003$, but not between Conditions 1 and 2, $t(16) = .35, p = .734$. As with regression path time, in sentences with the perfect structure, late bilinguals read the participle significantly more slowly when the switch occurred at the participle than when it occurred at the auxiliary. Nonetheless, they read the participles at a similar speed in both types of experimental sentences with the progressive structure.
4.4.3 Reading speed results

As mentioned in section 4.2.2, in addition to the Codeswitch block, participants also completed an English block and a Spanish block, each of which contained experimental sentences that had the same structure as the experimental codeswitched sentences as well as fillers. All the experimental sentences in both task blocks (i.e., the comprehension task and the acceptability judgment task blocks) were grammatically correct sentences. The fillers in the comprehension task block were also grammatically correct sentences. However, fillers in the acceptability judgment task block were ungrammatical sentences. Each sentence included one grammatical error, such as lack of subject-verb agreement, lack of gender or number agreement, or verb tense or mood errors. In order to calculate reading speed, the total time for each sentence was extracted. In the case of the experimental sentences, a two-way repeated measures ANOVA was conducted to evaluate the effect of structure and language on the extracted reading measure of total time. Structure (progressive versus perfect) and language (English versus Spanish) were the within-subjects factors. Separate analyses were conducted for each task and for each participant group.

Table 4-9 presents the mean total reading time of the experimental sentences by structure/language and participant group for the comprehension task. Standard deviations are displayed in parentheses.

<table>
<thead>
<tr>
<th>Structure/Language</th>
<th>Mean total reading time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early bilinguals</td>
</tr>
<tr>
<td>Progressive structure-English</td>
<td>5951.21 (SD = 2370.91)</td>
</tr>
<tr>
<td>Progressive structure-Spanish</td>
<td>8043.25 (SD = 3260.41)</td>
</tr>
<tr>
<td>Perfect structure-English</td>
<td>5983.84 (SD = 2075.32)</td>
</tr>
<tr>
<td>Perfect structure-Spanish</td>
<td>7809.95 (SD = 3269.10)</td>
</tr>
</tbody>
</table>
As seen in Table 4-9, the early bilinguals’ mean total reading time is faster for sentences in English with the progressive and perfect structures, when compared to sentences in Spanish. The results of the ANOVA displayed a main effect of language, $F(1,41) = 25.32, p < .001$, which shows that the difference in reading speed between the two languages is significant. The late bilinguals’ mean total reading time is slightly slower for the two types of Spanish sentences than for the two types of English sentences. However, the difference is not statistically significant, as demonstrated by the lack of a main effect of language in the ANOVA, $F(1,16) = 2.08, p = .169$.

Table 4-10 presents the mean total reading time of the filler sentences by language and by participant group for the comprehension task.

Table 4-10: Mean total reading time (in milliseconds) of the filler sentences by language and participant group for the comprehension task

<table>
<thead>
<tr>
<th>Language</th>
<th>Mean total reading time (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early bilinguals</td>
</tr>
<tr>
<td>Fillers-English</td>
<td>5658.39 (1943.27)</td>
</tr>
<tr>
<td>Fillers-Spanish</td>
<td>7444.86 (2620.52)</td>
</tr>
</tbody>
</table>

The same pattern that was found with the experimental sentences also appears with the filler sentences. In this case, paired-samples $t$ tests were conducted in order to examine differences between the total reading time means for each group of participants. The results for the early bilinguals indicated significant mean differences between English fillers and Spanish fillers, $t(41) = 5.43, p < .001$. Therefore, the Spanish fillers were read by these participants at a speed that was significantly slower, when compared to the speed at which they read the English fillers. On the other hand, the late bilinguals read both the English fillers and the Spanish fillers at a similar speed. The paired-samples $t$ test displayed no significant mean differences between both types of sentences for this group of participants, $t(16) = .81, p = .432$.

The mean total reading time of the experimental sentences by structure/language and participant group for the acceptability judgment task is displayed in Table 4-11.
The two-way repeated measures ANOVA for the early bilinguals presented a main effect of language, $F(1, 41) = 8.72$, $p = .005$. With the acceptability judgment task, the early bilinguals read English sentences at a significantly faster speed than the speed at which they read Spanish sentences. No main effect of language was found for the late bilinguals, $F(1, 16) = .31$, $p = .586$. When both the progressive and the perfect structures are taken together, these participants read English and Spanish sentences at a similar speed. Interestingly, the ANOVA for the late bilinguals displayed a main effect of structure, $F(1, 16) = 4.52$, $p = .050$, showing that sentences with the progressive structure were read significantly more quickly than sentences with the perfect structure, regardless of the language of the sentence. Table 4-12 exhibits the mean total reading time of the filler sentences by language and participant group for the acceptability judgment task. Recall that these fillers included a grammatical error somewhere in the sentence. This may explain why the total reading times for these fillers are slightly longer than those for the fillers in the comprehension task (Table 4-10).

Table 4-11: Mean total reading time (in milliseconds) of the experimental sentences by structure/language and participant group for the acceptability judgment task

<table>
<thead>
<tr>
<th>Structure/Language</th>
<th>Early bilinguals</th>
<th>Late bilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressive structure-English</td>
<td>7234.07 ($SD = 3048.66$)</td>
<td>7557.90 ($SD = 2735.39$)</td>
</tr>
<tr>
<td>Progressive structure-Spanish</td>
<td>8784.24 ($SD = 3460.57$)</td>
<td>7386.30 ($SD = 2427.68$)</td>
</tr>
<tr>
<td>Perfect structure-English</td>
<td>7540.16 ($SD = 3053.52$)</td>
<td>8298.74 ($SD = 3847.60$)</td>
</tr>
<tr>
<td>Perfect structure-Spanish</td>
<td>8848.61 ($SD = 3197.31$)</td>
<td>7821.42 ($SD = 2398.59$)</td>
</tr>
</tbody>
</table>

Table 4-12: Mean total reading time (in milliseconds) of the filler sentences by language and participant group for the acceptability judgment task

<table>
<thead>
<tr>
<th>Language</th>
<th>Mean total reading time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early bilinguals</td>
</tr>
<tr>
<td>Fillers-English</td>
<td>6591.50 ($SD = 2441.57$)</td>
</tr>
<tr>
<td>Fillers-Spanish</td>
<td>8182.41 ($SD = 2747.98$)</td>
</tr>
</tbody>
</table>
The paired-samples $t$ test that was conducted for the early bilinguals displayed significant mean reading time differences between the English fillers and the Spanish fillers, $t(41) = 3.94, p < .001$. Once again, these participants read the Spanish fillers at a significantly slower speed than the English fillers. The paired-samples $t$ test that was performed for the late bilinguals also indicated significant mean reading time differences between the English fillers and the Spanish fillers, $t(16) = 3.83, p = .001$. However, in this case, the effect went in the opposite direction. The late bilinguals read the Spanish fillers at a significantly faster speed than the English fillers.

4.5 Discussion

As with the State College study presented in Chapter 3, these results will be interpreted in relation to the three research questions. Table 4-13 presents a summary of the eye-tracking results by reading measure, participant group, and task.
The first research question asked whether the type of codeswitches that are more frequently found in natural production (i.e., *estar* + English participle switches) are easier to process during reading comprehension than the less frequent type of switches (i.e., *haber* + English participle switches).

Table 4-13 shows that, in the comprehension task, the results for both groups of participants displayed an interaction of auxiliary type and switch site in the measures of early processing (i.e., gaze duration and regression path time) and in the measure of later processing (i.e., total time).

For both early bilinguals and late bilinguals, the fixation durations on the participle were significantly longer for Condition 4 than for Condition 3. However, no significant differences were found between the fixation durations on the participle for Condition 1 and Condition 2. The fact that the participants in each group displayed the interaction in different measures of early processing – the early bilinguals displayed it in gaze duration and the late bilinguals displayed it in

### Table 4-13: Summary of the New York eye-tracking results by reading measure, participant group, and task

<table>
<thead>
<tr>
<th>(auxiliary type x switch site)</th>
<th>Early bilinguals</th>
<th>Late bilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comprehension task</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaze</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Total</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td><strong>Acceptability judgment task</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaze</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Total</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

√ Indicates a statistically significant effect \((p \leq .05)\).
regression path time—shows that they use different processing strategies when they face processing difficulties. The important finding is that both groups were sensitive to the asymmetrical occurrence of the \textit{estar}+English participle switches and the \textit{haber}+English participle switches in natural production. That is, from early on during their reading comprehension, it was easier for both groups of participants to process the more frequent \textit{estar}+English participle switches than the less frequent \textit{haber}+English participle switches, when each of these were compared to their corresponding conditions in which the switch occurred at the auxiliary.

The second research question examined the role of L2 acquisition during the comprehension of these codeswitches. When comparing the reading times of both groups of participants, overall, the late bilinguals’ fixation durations on the participle were longer than those of the early bilinguals (compare the means in Tables 4-4 and 4-7 for the early bilinguals with those in Tables 4-5 and 4-8 for the late bilinguals). Therefore, the codeswitched sentences may have been slightly harder to process for the late bilinguals than for the early bilinguals. There were also differences between the two groups regarding the particular effects displayed in Table 4-13, specifically in the case of the acceptability judgment task. As mentioned above, in the comprehension task, both groups of participants similarly displayed an interaction of auxiliary type and switch site in one of the two measures of early processing and in the measure of later processing. However, in the acceptability judgment task, each group of participants processed the codeswitched sentences differently. The early bilinguals only displayed a main effect of switch site during measures of early and later processing, showing that they had an easier time processing sentences that included a switch at the auxiliary, compared to sentences that included a switch at the participle. However, no by-participant interaction of auxiliary type and switch site emerged. Instead, the late bilinguals did exhibit a significant by-participant interaction of auxiliary type and switch site in the measure of regression path time and in the measure of total
time. Therefore, during early and later processing, it was harder for late bilinguals to process the past participle when it belonged to a sentence that included a switch at the participle, compared to when it belonged to a sentence that included a switch at the auxiliary. Nonetheless, there were no differences in processing difficulty for the late bilinguals when they read present participles that belonged to sentences with switches at the auxiliary or to those with switches at the participle. These findings lead to the conclusion that the two groups of participants display similarities during the comprehension task, but differences during the acceptability judgment task.

The third research question related to the way the different tasks influenced participants’ processing of the different types of codeswitches. Table 4-13 reveals clear task differences for the group of early bilinguals. In the comprehension task, these bilinguals exhibited a significant by-participant interaction of auxiliary type and switch site during early and later processing. The \textit{estar}+English participle switches were significantly easier to process than the \textit{haber}+English participle switches, when each of these were compared to their counterpart switches at the auxiliary. Therefore, when the early bilinguals had the task of reading for comprehension, they displayed sensitivity to the asymmetrical occurrence of these two types of codeswitches in natural production. In other words, when given a task that only requires comprehension of the codeswitched sentence, participants’ comprehension costs of these types of codeswitches reflected the production frequency of these codeswitches, as displayed in oral and written corpora. However, in the acceptability judgment task, the by-participant interaction of auxiliary type and switch site disappeared. When given this task, early bilinguals displayed more processing costs when reading participles that belong to sentences with a switch at the participle, regardless of the type of structure (i.e., progressive or perfect) to which the participle belonged. Thus, when the early bilinguals were asked to provide metalinguistic judgments of the codeswitched sentences, their processing costs no longer mirrored the frequency with which these types of codeswitches are naturally produced.
Contrary to the results of the early bilinguals, the findings for the group of late bilinguals did not reveal task effects. This group of bilinguals displayed an interaction of auxiliary type and switch site in regression path time and total time during both the comprehension task and the acceptability judgment task. The particular task given to the late bilinguals did not influence the way they processed the codeswitched sentences at all. In both cases, it was significantly easier for them to process the more frequent *estar*+English participle switches than the less common *haber*+English participle switches, when these were compared to their corresponding switches at the auxiliary. Therefore, it seems that the task given to participants does not influence the way all participants process codeswitched sentences.

Turning to participants’ actual judgments of the sentences, the results for the late bilinguals clearly coincided with their processing costs during the comprehension task and the acceptability judgment task. In both tasks, these participants had much more difficulty processing *haber*+English participle switches, when compared to perfect structures with a switch at the auxiliary. However, they processed *estar*+English participle switches and progressive structures with a switch at the auxiliary with the same ease. In their judgments, *haber*+English participle switches received the lowest percentage of acceptable judgments, specifically only 50%, while the other three types of switches received acceptable judgments over 90% of the time. The early bilinguals’ judgments coincided more so with the way they processed the different types of codeswitches during the comprehension task than during the acceptability judgment task. During the comprehension task, they had a harder time processing the *haber*+English participle switches than the perfect structures with a switch at the auxiliary. However, they processed *estar*+English participle switches similarly to the way they processed progressive structures with a switch at the auxiliary. In their judgments, the switches at the auxiliary and *estar*+English participle switches received acceptable judgments over 82% of the time. Nonetheless, *haber*+English participle switches were only judged acceptable approximately 64% of the time.
Based on the way these bilinguals processed the different types of switches during the acceptability judgment task, they would be expected to judge *estar*+English participle switches and *haber*+English participle switches in the same way, but that was not the case.

Finally, the answers to the comprehension questions do not reflect the differences that both groups of participants presented when processing the *estar*+English participle switches and the *haber*+English participle switches in the comprehension task. In this task, although it was easier for all participants to process the *estar*+English participle switches than the *haber*+English participle switches, the comprehension questions to all types of sentences were answered correctly between 87% and 93% of the time. Therefore, the processing difficulties with *haber*+English participle switches were captured by the eye-tracking reading measures, but they were overcome by the time the participants had to answer the comprehension questions and they did not impede the overall comprehension of the codeswitched sentences.

The reading speed results showed that the early bilinguals were consistently faster when reading sentences in English than sentences in Spanish. This difference was significant with both the experimental sentences and the filler sentences. The significant difference between reading speed in English and Spanish also emerged for the early bilinguals regardless of the task that they were given. This result is not surprising given that the early bilinguals, although proficient in both languages, are surely more dominant in English. On the other hand, in most cases, the late bilinguals displayed no significant differences between their reading speed in English and in Spanish. Although they acquired English as a later L2 and moved to the U.S. later in life, they are now students in a U.S. college and have increased exposure to English in oral and written language. This experience has undoubtedly helped them increase their reading speed in English. The one exception is the case of the filler sentences in the acceptability judgment task, with which the late bilinguals displayed significant mean reading time differences between the English
fillers and the Spanish fillers. They were significantly faster when reading the Spanish fillers than the English fillers.

4.6 Conclusion

This study re-examined the same research questions presented in Chapter 3. That is, this study was aimed at identifying whether more frequently produced types of codeswitches were easier to process than less frequently produced types of codeswitches. The other objectives were to study the influence of age of L2 acquisition and of task on the way Spanish-English bilinguals processed frequent and infrequent types of codeswitches. The results indicated that frequent codeswitches were easier to process than less frequent codeswitches by both early bilinguals and late bilinguals. One can assume that these participants are exposed to the estar+English participle switches more often than to the haber+English participle switches both in oral and written language. Therefore, when they are faced with these two types of switches, it is easier for them to process the former, which they have encountered more frequently in the past, than the latter. These results are consistent with constraint satisfaction models of language processing, such as the Production-Distribution-Comprehension approach (Gennari & MacDonald, 2009, MacDonald & Thornton, 2009), that attribute a great deal of importance to previous experience with linguistic structures when it comes to the way that individuals process language.

The findings also exhibited differences based on task and participant group. Specifically, the results described above were only found for the group of early bilinguals in the comprehension task, but not in the acceptability judgment task. When these participants were asked to read the codeswitched sentences in order to judge them, the processing costs they experienced with estar+English participle switches were similar to those that they experienced with haber+English participle switches, when each of these were compared with baseline
codeswitched conditions in which the switch occurred at the auxiliary. However, the group of late bilinguals displayed the same processing patterns of these two types of codeswitches in the comprehension task and in the acceptability judgment task. Therefore, the particular task given to participants affected the processing costs associated with codeswitched sentences, at least in the case of bilinguals who have been exposed to codeswitching for larger portions of their life.
Chapter 5

Switch costs

Numerous researchers have focused on the phenomenon of alternating languages in order to examine the presence of switch costs during bilingual production and comprehension. Several psycholinguistic studies have focused on controlled language production. In a seminal study on switch costs, Macnamara, Krauthammer, and Bolgar (1968) presented three groups of participants (French-speaking learners of English, English-speaking learners of French, and balanced French-English bilinguals) with a series of slides. Each slide included a number (from 1 to 20) inside a particular geometric shape (a circle, a triangle, a square, or a diamond). Participants’ task was to either name the number presented on the slide or to add “1” to the presented number and name the result. In addition, participants were asked to complete the task in English or French. The researchers controlled the delay between participants’ response and the presentation of the following slide; there was either no delay or a 2-second delay. Participants completed unilingual blocks and switching blocks. In the switching blocks, the researchers also controlled the pattern of switches, which was either regular (predictable) or irregular (unpredictable). The geometric shape surrounding the number indicated to participants the particular type of response that was required and participants responded verbally. Analyses showed that for all three groups of participants the requirement to switch languages added significantly to their response latencies. This was the case in lists with both regular and irregular switching patterns, although in lists with irregular switching patterns the increase in response latencies was even larger (Macnamara et al., 1968, p. 211). Participants’ response latencies when they had to switch languages were very similar to their response latencies when they had to switch numbers (i.e., name the presented
number plus “1”). These findings led the researchers to conclude that language switching and numerical switching involved similar skills (p. 212).

A similar study was conducted by Meuter and Allport (1999), in which bilinguals of English and another language completed the digit naming task. In this case, the participants were presented with a number (from 1 to 9) that was surrounded by a blue or a yellow rectangle, which cued the language in which they should name the number (either in their first language [L1] or their second language [L2]). Although the color-cue presentation was not completely predictable, on every trial there was a 0.3 probability of a switch. Each stimulus appeared on the computer screen 400 milliseconds (ms) after the participants triggered the voice key with their response to the previous trial. Switch trials were trials in which the background color was different from that of the preceding trial and nonswitch trials were trials in which the background color was the same as that of the preceding trial. Response latency results showed that language switching involved a processing cost, but this cost was larger when participants switched to their dominant L1 than when they switched to their weaker L2. The researchers also found that relative proficiency in both languages played a role in the language switch cost. When they divided the participants into two groups (a group with lower L2 proficiency and, therefore, a larger difference between L1 and L2 proficiency, and a group with higher L2 proficiency and, therefore, less difference between L1 and L2 proficiency), they found that the switch cost asymmetry decreased in the case of the more balanced bilinguals (Meuter & Allport, 1999, p. 35).

Some studies have replicated the asymmetrical switch cost (e.g., Costa, Santesteban, & Ivanova, 2006; Jackson, Swainson, Cunningham, & Jackson, 2001), while others have found evidence that constrains or reduces its generality. For example, Finkbeiner, Almeida, Janssen, and Caramazza (2006) compared L1 switch costs for univalent stimuli (stimuli that elicited responses exclusively in participants’ L1) and bivalent stimuli (stimuli that elicited both L1 and L2 responses). The participants were native English speakers with relative proficiency in a
variety of second languages. The researchers presented participants with the same numeral naming task that Meuter and Allport (1999) did, where they had to name the digit in their L1 or L2 depending on the background color (bivalent stimuli). However, Finkbeiner et al. also included two different types of univalent stimuli, either pictures or dot patterns, which participants only named in their L1. Switch costs were found when participants named bivalent stimuli, but not when they named univalent stimuli. Moreover, Finkbeiner et al. found an asymmetrical switch cost in a unilingual task that manipulated the speed of word availability. Participants were presented with univalent stimuli and asked to name either the ink color of the word presented or the word itself. In some cases, the words were easily accessible (e.g., *house*) and, in other cases, they were harder to access (e.g., *cottage*). The researchers concluded that the processes that contribute to switch costs are sensitive to the specific properties of individual lexical items (e.g., speed of response availability) rather than relative strength of bilinguals’ two languages (Finkbeiner et al., 2006, p. 1085).

Costa and Santesteban (2004) conducted a study in which participants were asked to name pictures of common objects, rather than Arabic numerals. The response language was signaled by the color of the picture (either red or blue). Each trial was presented with an interval of 1150 ms from the previous trial. As in Meuter and Allport’s study, 30% of the trials constituted switch trials and the remaining 70% constituted nonswitch trials. The asymmetrical switch cost reported by Meuter and Allport (1999) was replicated by Costa and Santesteban (2004), but only with the groups of unbalanced L2 learners (Spanish-Catalan and Korean-Spanish). For these participants, switching from their L2 to their L1 was harder than vice versa. However, with highly proficient Spanish-Catalan bilinguals, the magnitude of the switch cost was the same for the L1 and the L2. Moreover, the participants presented a reversal of dominance, exhibited by the fact that their naming latencies were faster in the L2 than in the L1. The
symmetrical switch costs were even present when the bilinguals performed the task in their dominant L1 and their much weaker third language (L3).

Costa and Santesteban (2004) conducted an additional experiment in which they altered the presentation of the pictured objects in two ways. Instead of simultaneous presentation of the object and its language cue, each picture was preceded by the presentation of the language cue, which was displayed for 300 ms. In addition, the highly proficient bilinguals were divided into two groups, and each received a different delay before the picture presentation. For one group, the picture appeared 500 ms after the language cue onset whereas for the other group each picture appeared 800 ms after the cue onset. Findings showed that the symmetrical switch costs persisted, that is, the difference between L1 and L2 naming latencies was unaffected by the time given to participants to prepare the response language (p. 504).

Other studies in which highly proficient bilinguals complete a picture-naming task also find evidence of symmetrical switch costs (e.g., Abutalebi et al., 2007). A study conducted by Gollan and Ferreira (2009) even found symmetrical switch costs with unbalanced bilinguals. In this case, Spanish-English bilinguals were asked to name images in three conditions: in their dominant language, in their weaker language, and in the first language that came to mind. The researchers found that, when unbalanced bilinguals were allowed to manage the switching themselves, they exhibited symmetrical switch costs, similarly to those found with balanced bilinguals in cued switching studies (e.g., Costa & Santesteban, 2004). Moreover, the researchers found that unlike cued switching, voluntary switching sometimes facilitated responses. For instance, English-dominant bilinguals named pictures in Spanish more quickly in the either-language condition than in the Spanish-only condition (p. 648).

Whereas most studies have examined switching costs associated with the production of isolated lexical items (nouns or numerals), a recent study by Tarlowski, Wodniecka, & Marzecová (2012) focused on the production of verbal grammatical structures. In this study, a
group of Polish-English unbalanced bilinguals were presented with pictures depicting completed actions or actions in progress. They were asked to describe the picture in either their L1 (Polish) or their L2 (English) according to an auditory cue that was presented in one of the two languages. Participants completed one block of pictures showing ongoing action, in which the progressive structure was elicited, and another block of pictures showing completed actions, in which the perfect structure was elicited. Within each block, participants completed switch trials and nonswitch trials. The researchers found different results depending on the grammatical structure that was elicited. In the block that elicited the perfect structure, participants presented asymmetrical switch costs, in which switching to the L1 was more costly than switching to the L2, consistent with Meuter and Allport’s (1999) findings. However, in the block that elicited the progressive structure, participants exhibited symmetrical switch costs accompanied by reverse dominance, in which participants displayed faster naming latencies in their L2 than in their L1 (p. 12). These results show that switching of unconnected lexical items differs from switching of phrases. In other words, the nature of switch costs varies depending on the characteristics of the particular utterances that bilinguals produce.

Thus far, these studies have shown that language switching costs, either symmetrical or asymmetrical, emerge when balanced and unbalanced bilinguals take part in controlled production studies. Contradictory evidence was found in a study that examined switch costs in a more naturalistic environment with free production. Timm (1983) measured the pauses of a portion of Spanish-English codeswitched speech produced by a Chicano adult male disc-jockey taped in 1979 on the KPFA radio station. The program was aimed at bilingual Chicano young adults in the San Francisco Bay Area, thus, the speaker used youth-oriented, popular varieties of both languages. Timm found that similar amounts of pauses occurred in unilingual discourse and in codeswitched discourse. In addition, the pauses in unilingual discourse and codeswitched
discourse were of equal length (p. 409). These results led to the conclusion that, during free
production, codeswitched speech is not more costly than unilingual speech.

Fewer studies have examined switch costs during the comprehension of codeswitches.
Altarriba, Kroll, Sholl, and Rayner (1996) used the eye-tracking technique to examine the degree
to which sentence context effects operate at a lexical or a conceptual level. Highly proficient
Spanish-English bilinguals read English sentences that included a target word (a concrete noun)
towards the middle of the sentence. That target word was either in English (unswitched
condition) or in Spanish (switched condition). The researchers controlled lexical frequency of the
target words (either high- or low-frequency) as well as sentence constraint (either high- or low-
constraint). Although the main objective of the study was not to examine switch costs, the
researchers reported consistent differences in the bilinguals’ processing of the English
(unswitched) and the Spanish (switched) target words. When the target words were unswitched,
participants were more likely to skip them. In cases where the target words were fixated, first
fixation durations and gaze durations were shorter for English target words than for Spanish
target words (p. 481). The researchers also conducted a study with the same sentences, but using
the rapid serial visual presentation (RSVP) method, in which participants read the sentences one
word at a time and were asked to name aloud the capitalized target word. The results for mean
naming latencies on the target words also showed switch costs; participants took longer to name
the switched target words than the unswitched target words (p. 484).

Two recent electrophysiological studies have also examined participants’ comprehension
of single-word switches. In these studies, participants read sentences one word at a time. These
sentences appeared in one language and included a codeswitch into the other language at the final
word of the sentence (generally a noun). Proverbio, Leoni, and Zani (2004) used a blocked
language design, therefore allowing participants to know when they would encounter final-word
switches. In this case, Italian-English simultaneous interpreters read the sentences in separate
blocks: unilingual Italian, unilingual English, Italian ending with an English word, and English ending with an Italian word. The results displayed switch costs during sentence comprehension, exhibited by larger N400 effects in codeswitched conditions, despite the fact that language presentation was blocked. In addition, participants displayed a larger N400 effect when codeswitched sentences ended with words in their L2 (English) than when they ended with words in their L1 (Italian). Moreover, the interpreters were faster at reading and comprehending sentences in English that ended with an Italian word than vice versa (Proverbio et al., 2004, p. 1647). This asymmetrical switch cost appears to go in the opposite direction from that which has been found in the production studies discussed above. In this case, it was more costly for the bilinguals to process sentences that switched into the L2 compared to those that switched into the L1.

Moreno, Federmeier, and Kutas (2002) conducted a similar study with English-Spanish highly proficient bilinguals. In this case, the researchers used a mixed presentation design. The sentences that participants read always began in English, but could end with three different types of target words: an expected completion, a congruent, but low probability completion (referred to as a lexical switch), or a literal Spanish translation of the expected completion (referred to as a codeswitch). The results of this study also displayed switch costs, but with a different event-related potential (ERP) component. Participants exhibited larger Late Positive Complex (LPC) effects with sentences that ended in a codeswitch compared to those that ended with the expected completion (p. 196). Moreno et al. also found that the switch cost was modulated by level of proficiency in the bilinguals’ L2. Bilinguals who were more balanced (who were also more likely to report frequent codeswitching in everyday life) displayed less of an LPC response to the final-word codeswitches than the less balanced bilinguals, showing that the switch was less unexpected or less difficult for them to process (p. 202). Moreover, the researchers found that codeswitches and lexical switches elicited qualitatively different ERP responses. Lexical switches elicited an
increased N400 response, suggesting that they were more difficult to integrate semantically with the sentence context than the expected completion. However, the negative response elicited by the codeswitches differed in its distributional characteristics from an N400. This finding led the researchers to conclude that processing a codeswitch produces less of a lexical-semantic processing cost than a within-language lexical switch (p. 203).

The present study adds to the body of research that examines the switch costs associated with the comprehension of codeswitches. Unlike previous studies that have examined switch costs in comprehension, this study examines the way that bilinguals process, not single-word switches, but intrasentential switches in which entire phrases from both languages appear within a single sentence. Specifically, this experiment was aimed at investigating whether it was more costly for bilinguals to read and comprehend codeswitched sentences than unilingual sentences.

5.1 Method

5.1.1 Participants

Because this study was conducted as part of the main eye-tracking study described in Chapter 4, the participants who completed that study also completed the one described here. Therefore, two groups of Spanish-English bilinguals completed this study: a group of early bilinguals and a group of late bilinguals. Both groups were proficient in both languages and reported using and being exposed to both languages on a regular basis. The early bilinguals arrived in the United States (U.S.) at a younger age than the late bilinguals. In addition, the early bilinguals acquired both Spanish and English at similar ages whereas the late bilinguals acquired their second language (English) at a later age. Overall, the early bilinguals proved to be more proficient in English while the late bilinguals demonstrated more proficiency in Spanish (as
shown by the measures of language proficiency used in the study that was described in Chapter 4).

5.1.2 Materials and Design

As mentioned in Chapter 4, a series of unilingual English sentences were added to the eye-tracking experiment that examined the comprehension of codeswitches. The experimental English sentences constituted direct translations of the codeswitched experimental sentences. Therefore, two additional versions were added to the 96 original codeswitched item sets (Appendix D). These two versions were unilingual English sentences, one with the progressive structure and one with the perfect structure (Appendix E). In order to examine switch costs, the two unilingual English versions of the sentences were compared with the two codeswitched versions of the sentences in which the switch occurred at the auxiliary (labeled Conditions 1 and 3 in Chapter 4). These four conditions are displayed in Table 5-1, and the critical region examined here is underlined. The conditions have been renumbered in this chapter. Conditions 1 and 3 from Chapter 4 are Conditions 2 and 4 in this chapter.

Table 5-1: Sample sentences used to examine switch costs

<table>
<thead>
<tr>
<th>Condition</th>
<th>Sample sentence</th>
</tr>
</thead>
</table>
| (1) Progressive-Unilingual English | The director confirmed that the actors 
are rehearsing their lines for the movie. |
| (2) Progressive-Switch at auxiliary | El director confirmó que los actores 
are rehearsing their lines for the movie. |
| (3) Perfect-Unilingual English | The director confirmed that the actors 
have rehearsed their lines for the movie. |
| (4) Perfect-Switch at auxiliary | El director confirmó que los actores 
have rehearsed their lines for the movie. |

The difference between Conditions 1 and 3 and Conditions 2 and 4 is that the beginning of the sentence up to the point of the codeswitch is in Spanish in Conditions 2 and 4. The mean
character length of the first noun in the experimental sentences was seven (range 3-14) for the English words and eight (range 4-13) for the Spanish words. The mean lexical frequency of the first noun was 123.62 (range 2-994) for the English words and 45.39 (range 1-173) for the Spanish words. A Spanish noun used in one sentence exceeded that range: the word *hombre* ‘man,’ which had a frequency of 1965. The mean character length of the first verb in the experimental sentences was seven (range 4-10) for the English words and five (range 4-9) for the Spanish words. The mean lexical frequency of the first verb was 88.38 (range 2-866) for the English words and 78.61 (range 1-177) for the Spanish words. One verb in both languages exceeded those ranges: the English verb *said*, which had a frequency of 3922, and the Spanish verb *dijo* ‘said,’ which had a frequency of 1607. The mean character length of the second noun in the experimental sentences, that is, the noun in the subordinate clause was nine (range 5-13) for the English words and nine (range 5-14) for the Spanish words. The mean lexical frequency of the second noun was 54.78 (range 2-426) for the English words and 37.37 (range 1-131) for the Spanish words. Like the codeswitched sentences, the unilingual English sentences were divided into six reading files, each including 16 experimental sentences (eight sentences from Condition 1 and eight sentences from Condition 3) as well as 16 filler sentences, described in section 4.2.2 of the previous chapter. The design included a blocked language presentation of the sentences. Therefore, participants read the codeswitched sentences in one block and the unilingual English sentences in another block. Within each block, participants completed the two tasks described in Chapter 4: a comprehension task and an acceptability judgment task. Language block

---

8 Lexical frequencies for Spanish words are from the Alameda and Cuetos (1995) corpus and were obtained through the *Normas e índices de interés en Psicología Experimental* website (Díez, Fernández, & Alonso, 2006). The frequencies for English words are from the Kučera and Francis (1967) corpus and were obtained through the *The English Lexicon Project* website (Balota et al., 2007). Because the Spanish corpus is based on two million words while the English corpus is based on one million words, for purposes of a comparison between them, the English word frequencies were multiplied by 2.
presentation and task block presentation order was counterbalanced across participants, and the sentences within each block were pseudo-randomized.

5.1.3 Procedure

Because this experiment was part of the experiment described in Chapter 4, the procedure is identical to that described in section 4.2.3. Participants spent approximately 10-15 minutes to complete each of the unilingual English blocks and 20-25 minutes to complete each of the Codeswitched blocks.

5.2 Results

The results were extracted and analyzed in the same way as described in Chapter 4, section 4.4. However, there was one crucial difference. In this case, the critical region under analysis comprised the auxiliary in the subordinate clause of each experimental sentence (underlined in Table 5-1 for ease of exposition). In the case of the codeswitched sentences (Conditions 2 and 4 in Table 5-1), this was the point at which the codeswitch from Spanish into English occurred. Therefore, if the participants experienced switch costs during reading comprehension, it was expected to surface initially at this point. The two measures of early processing (i.e., gaze duration and regression path time) and the measure of later processing (i.e., total time) were extracted for the auxiliary in the experimental sentences. Analyses were conducted separately for each task and for each participant group. The results for the comprehension task will be presented first, followed by those for the acceptability judgment task.
5.2.1 Results for the comprehension task

5.2.1.1 Early bilinguals

Table 5-2 presents the mean gaze duration, regression path time, and total time by condition for the early bilinguals during the comprehension task. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-Unilingual English</td>
<td>131.77 (SD = 86.19)</td>
<td>175.38 (SD = 146.21)</td>
<td>254.05 (SD = 178.76)</td>
</tr>
<tr>
<td>(2) Progressive-Switch at auxiliary</td>
<td>136.60 (SD = 93.21)</td>
<td>174.70 (SD = 151.47)</td>
<td>293.96 (SD = 168.63)</td>
</tr>
<tr>
<td>(3) Perfect-Unilingual English</td>
<td>182.22 (SD = 96.15)</td>
<td>215.41 (SD = 124.37)</td>
<td>362.70 (SD = 189.59)</td>
</tr>
<tr>
<td>(4) Perfect-Switch at auxiliary</td>
<td>187.53 (SD = 100.75)</td>
<td>261.38 (SD = 172.67)</td>
<td>381.10 (SD = 197.49)</td>
</tr>
</tbody>
</table>

A two-way repeated measures analysis of variance (ANOVA) was performed in order to evaluate the effect of switch and auxiliary type on gaze duration, regression path time, and total time. Switch (switch versus nonswitch) and auxiliary type (progressive versus perfect) were the within-subjects factors.

Gaze duration. For gaze, the results did not yield a main effect of switch, $F(1,41) = .15, p = .701$, nor a by-participant interaction of switch and auxiliary type, $F(1,41) = .00, p = .975$. There was only a main effect of auxiliary type, $F(1,41) = 27.16, p < .001$, such that the progressive auxiliary was read more quickly than the perfect auxiliary. Therefore, the early bilinguals did not exhibit any gaze duration differences between the auxiliaries that belonged to unilingual sentences and those than belonged to codeswitched sentences.
Regression path time. For regression path time, the same results arose. There was no main effect of switch, $F(1,41) = 1.37, p = .249$, and no by-participant interaction of switch and auxiliary type, $F(1,41) = 2.28, p = .139$. However, a main effect of auxiliary type was found, $F(1,41) = 17.76, p < .001$. Regression path reading times on the auxiliary displayed no switch costs for the early bilinguals. Again, the progressive auxiliary was read more quickly than the perfect auxiliary.

Total time. For total time, once again the same results emerged for the early bilinguals: no main effect of switch, $F(1,41) = .97, p = .330$ and no by-participant interaction of switch and auxiliary type, $F(1,41) = .57, p = .453$. There was, however, a main effect of auxiliary type, $F(1,41) = 43.72, p < .001$. Early bilinguals consistently read the progressive auxiliary more quickly than the perfect auxiliary during the comprehension task. Once again, they did not display a switch cost.

5.2.1.2 Late bilinguals

Table 5-3 includes the mean gaze duration, regression path time, and total time by condition for the late bilinguals during the comprehension task, with standard deviations shown in parentheses. These participants exhibited the same pattern of results as the early bilinguals during this task.
Gaze duration. The results for gaze duration failed to show a main effect of switch, $F(1,16) = .43, p = .520$, as well as an interaction of switch and auxiliary type, $F(1,16) = .79, p = .387$. Therefore, no switch costs were observed. Nonetheless, a main effect of auxiliary type emerged, $F(1,16) = 23.45, p < .001$, showing that the progressive auxiliary was read by the late bilinguals significantly more quickly than the perfect auxiliary.

Regression path time. Regression path time results also brought about no main effect of switch, $F(1,16) = .84, p = .374$, and no by-participant interaction of switch and auxiliary type, $F(1,16) = 1.01, p = .329$. A main effect of auxiliary type was present, $F(1,16) = 10.85, p = .005$. Again, the late bilinguals were faster when reading the progressive auxiliary than the perfect auxiliary, but they did not display any switch costs.

Total time. For total time, the late bilinguals’ results did not generate a main effect of switch, $F(1,16) = .00, p = .997$, nor a by-participant interaction of switch and auxiliary type, $F(1,16) = .62, p = .441$, but a main effect of auxiliary type did arise, $F(1,16) = 23.07, p < .001$. The late bilinguals’ results were consistent across the three measures of processing. They always read the progressive auxiliary more quickly than the perfect auxiliary. However, they never displayed significantly different reading times between the auxiliaries in the unilingual sentences and those in the codeswitched sentences.

Table 5-3: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the late bilinguals during the comprehension task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-</td>
<td>122.03</td>
<td>156.07</td>
<td>272.70</td>
</tr>
<tr>
<td>Unilingual English</td>
<td>$(SD = 114.94)$</td>
<td>$(SD = 126.85)$</td>
<td>$(SD = 196.46)$</td>
</tr>
<tr>
<td>(2) Progressive-</td>
<td>120.01</td>
<td>146.89</td>
<td>255.71</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>$(SD = 111.96)$</td>
<td>$(SD = 174.60)$</td>
<td>$(SD = 184.22)$</td>
</tr>
<tr>
<td>(3) Perfect-</td>
<td>166.19</td>
<td>262.46</td>
<td>391.25</td>
</tr>
<tr>
<td>Unilingual English</td>
<td>$(SD = 89.98)$</td>
<td>$(SD = 158.95)$</td>
<td>$(SD = 284.49)$</td>
</tr>
<tr>
<td>(4) Perfect-</td>
<td>188.71</td>
<td>214.35</td>
<td>407.90</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>$(SD = 94.27)$</td>
<td>$(SD = 157.26)$</td>
<td>$(SD = 193.79)$</td>
</tr>
</tbody>
</table>
5.2.2 Results for the acceptability judgment task

5.2.2.1 Early bilinguals

Table 5-4 displays the mean gaze duration, regression path time, and total time by condition for the early bilinguals in the acceptability judgment task, with standard deviations shown in parentheses.

Table **5-4**: Mean gaze duration, regression path time, and total time (in milliseconds) by condition for the early bilinguals during the acceptability judgment task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-</td>
<td>138.71</td>
<td>186.54</td>
<td>419.07</td>
</tr>
<tr>
<td>Unilingual English</td>
<td>$(SD = 90.44)$</td>
<td>$(SD = 157.97)$</td>
<td>$(SD = 296.98)$</td>
</tr>
<tr>
<td>(2) Progressive-</td>
<td>203.27</td>
<td>269.28</td>
<td>455.51</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>$(SD = 114.68)$</td>
<td>$(SD = 164.56)$</td>
<td>$(SD = 227.57)$</td>
</tr>
<tr>
<td>(3) Perfect-</td>
<td>198.83</td>
<td>250.57</td>
<td>619.57</td>
</tr>
<tr>
<td>Unilingual English</td>
<td>$(SD = 99.09)$</td>
<td>$(SD = 149.84)$</td>
<td>$(SD = 357.97)$</td>
</tr>
<tr>
<td>(4) Perfect-</td>
<td>235.33</td>
<td>302.21</td>
<td>622.43</td>
</tr>
<tr>
<td>Switch at auxiliary</td>
<td>$(SD = 114.66)$</td>
<td>$(SD = 135.66)$</td>
<td>$(SD = 384.71)$</td>
</tr>
</tbody>
</table>

Another 2 x 2 ANOVA evaluated the effect of switch and auxiliary type on gaze duration, regression path time, and total time, with switch (switch versus nonswitch) and auxiliary type (progressive versus perfect) and as within-subjects factors.

**Gaze duration.** In the measure of gaze duration, the early bilinguals displayed a main effect of switch, $F(1,41) = 9.64, p = .003$, as well as a main effect of auxiliary type, $F(1,41) = 14.98, p < .001$. No by-participant interaction of switch and auxiliary type was found, $F(1,41) = 1.35, p = .252$. Therefore, with the acceptability judgment task, the group of early bilinguals displayed longer gaze durations on the auxiliaries when they were included in codeswitched sentences, compared to when they were in unilingual sentences. Also, reading progressive auxiliaries took less time than reading perfect auxiliaries.
Regression path time. Similar results emerged for the measure of regression path time: a main effect of switch, $F(1,41) = 7.12, p = .011$, a main effect of auxiliary type, $F(1,41) = 7.31, p = .010$, but no by-participant interaction of switch and auxiliary type, $F(1,41) = .57, p = .456$. Again, the early bilinguals took longer to read auxiliaries in codeswitched sentences than those in unilingual sentences. They also took longer to read the perfect auxiliary than the progressive auxiliary.

Total time. For total time, the early bilinguals’ results only displayed a main effect of auxiliary type, $F(1,41) = 19.39, p < .001$. There was no main effect of switch, $F(1,41) = .15, p = .698$, and no by-participant interaction of switch and auxiliary type, $F(1,41) = .28, p = .600$. In this measure of later processing, the early bilinguals no longer presented reading time differences between the auxiliaries in unilingual sentences and those in codeswitched sentences. They did, however, read progressive auxiliaries more quickly than perfect auxiliaries.

5.2.2.2 Late bilinguals

Table 5-5 displays the mean gaze duration, regression path time, and total time by condition for the late bilinguals during the acceptability judgment task.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gaze duration</th>
<th>Regression path time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Progressive-Unilingual English</td>
<td>149.10 ($SD = 143.18$)</td>
<td>201.15 ($SD = 209.74$)</td>
<td>324.89 ($SD = 259.73$)</td>
</tr>
<tr>
<td>(2) Progressive-Switch at auxiliary</td>
<td>173.73 ($SD = 204.20$)</td>
<td>245.01 ($SD = 323.05$)</td>
<td>432.09 ($SD = 310.01$)</td>
</tr>
<tr>
<td>(3) Perfect-Unilingual English</td>
<td>203.64 ($SD = 125.01$)</td>
<td>287.75 ($SD = 204.85$)</td>
<td>672.03 ($SD = 378.58$)</td>
</tr>
<tr>
<td>(4) Perfect-Switch at auxiliary</td>
<td>183.89 ($SD = 139.70$)</td>
<td>260.83 ($SD = 250.10$)</td>
<td>424.24 ($SD = 286.53$)</td>
</tr>
</tbody>
</table>
**Gaze duration.** For gaze duration, the late bilinguals presented no main effect of switch, $F(1,16) = .01, p = .938$, no main effect of auxiliary type, $F(1,16) = 1.34, p = .265$, and no by-participant interaction of switch and auxiliary type, $F(1,16) = 1.31, p = .268$. Therefore, the auxiliaries in all four experimental conditions were read at a similar speed during this measure of early processing.

**Regression path time.** The same lack of effects was displayed in the measure of regression path time. There was no main effect of switch, $F(1,16) = .03, p = .869$, no main effect of auxiliary type, $F(1,16) = 3.25, p = .091$, and no by-participant interaction of switch and auxiliary type, $F(1,16) = .88, p = .364$. In this measure as well, the late bilinguals did not present auxiliary reading time differences in any of the four experimental conditions.

**Total time.** For total time, the late bilinguals’ results did not yield a main effect of switch, $F(1,16) = 3.33, p = .087$. However, there was a main effect of auxiliary type, $F(1,16) = 11.54, p = .004$, as well as a significant by-participant interaction of switch and auxiliary type, $F(1,16) = 19.87, p < .001$. Follow-up paired-samples $t$ tests displayed significant mean differences between Conditions 1 and 2, $t(16) = 3.78, p = .002$, and between Conditions 3 and 4, $t(16) = 3.39, p = .004$. During the acceptability judgment task, the late bilinguals read progressive auxiliaries more quickly when they belonged to unilingual sentences than when they belonged to codeswitched sentences. The opposite was true for the perfect auxiliaries; they were read more quickly when they belonged to codeswitched sentences.

### 5.3 Discussion

This study set out to examine the presence of switch costs for two groups of Spanish-English bilinguals (early bilinguals and late bilinguals) while they read codeswitched sentences. Two types of sentences, which were presented in separate language blocks, were compared:
unilingual English sentences and sentences that started in Spanish and switched to English. Within each language block, some sentences included the progressive structure and others included the perfect structure. Moreover, participants were asked to read these sentences in order to complete two different tasks. In one task block, they were asked to read the sentences to answer comprehension questions about them. In the other, they were asked to read the sentences to perform an acceptability judgment on them.

In order to measure switch costs, reading times on the auxiliary of the progressive and perfect structures were calculated because this was the point at which the codeswitch occurred in the mixed sentences. The results for the comprehension task were consistent across both groups of participants. When reading for comprehension, reading times at the auxiliary in the unilingual sentences were similar to reading times at the auxiliary in the codeswitched sentences. This result suggests that participants experienced no switch costs when they were asked to read for comprehension. Another result that was consistent across both groups of participants during the comprehension task was that the progressive auxiliaries were read more quickly than the perfect auxiliaries. This result may have to do with the form of each auxiliary. In these stimuli, the auxiliaries were always in their plural form, that is, *are* in the case of the progressive auxiliary and *have* in the case of the perfect auxiliary. It is possible that the longer reading times for the perfect auxiliaries are simply due to the fact that *have* is longer than *are*. Another possibility is that the progressive *are* is more frequent than the perfect *have*. Past research has shown that early measures of processing (i.e., gaze duration and regression path) are sensitive to lexical characteristics of words, such as length and frequency (Juhasz & Pollatsek, 2011). It is possible, then, that the main effect of auxiliary type reflects the different lexical characteristics of the two types of auxiliaries. The crucial finding for this study, however, was that no switch costs were found for either of the two participant groups with the comprehension task.
The results for the acceptability judgment task differed from the comprehension task results. They also differed between participant groups. The only similarity between the comprehension task results and the acceptability judgment task results was that the early bilinguals, once again, displayed longer reading times for the perfect auxiliaries than for the progressive auxiliaries. However, the early bilinguals exhibited differences between their auxiliary reading times in the two tasks in that the auxiliary fixation durations were longer when they appeared in codeswitched sentences than when they appeared in unilingual sentences. This result was exhibited in the measures of gaze duration and regression path time, but not in the measure of total time. Therefore, when they were asked to read the sentences in order to perform metalinguistic judgments, the early bilinguals displayed switch costs in the measures of early processing. The switch costs disappeared, however, in the measure of later processing. It is evident that the task, at least temporarily, altered the early bilinguals’ processing in a way that led them to exhibit switch costs when reading codeswitched sentences.

The late bilinguals displayed a different pattern of reading times during the acceptability judgment task. In the measures of early processing, there were no differences between the auxiliary reading times for any of the four conditions. Thus, progressive and perfect auxiliaries were read at similar speeds in both unilingual and codeswitched sentences. However, in the measure of later processing, the late bilinguals displayed reading time differences between all four conditions. Progressive auxiliaries were read more quickly in unilingual sentences than in codeswitched sentences, but perfect auxiliaries were read more quickly in codeswitched sentences than in unilingual sentences. These results suggest that the late bilinguals displayed switch costs with the progressive structures, but not with the perfect structures. The large processing cost that late bilinguals exhibited when reading perfect auxiliaries in unilingual sentences is difficult to explain. It is possibly just an effect of the acceptability judgment task.
In conclusion, these results suggest that task strongly influences the way bilinguals process codeswitched sentences. When bilinguals are asked to read sentences in order to judge their acceptability, they display switch costs, at least with some linguistic structures and for some reading measures. However, when the task only entails reading for comprehension, it is not more difficult for bilinguals to read codeswitched sentences than it is for them to read unilingual sentences. Although this comprehension task is experimental, between both tasks, it is the one that most closely resembles what individuals experience during normal, free reading. These results coincide with findings for free codeswitched production by Timm (1983) and Gollan and Ferreira (2009) and, therefore, display a correspondence between production and comprehension of codeswitches. When bilingual speakers are free to codeswitch wherever they want, they do not exhibit switch costs. During comprehension, the codeswitches are unexpected and can potentially be more difficult to process, particularly in the case of written language, in which codeswitching is less common. However, these results show that, when bilingual codeswitchers read with the sole purpose of understanding what they are reading, they do not display switch costs.

There are two caveats to consider. First, in this experimental design, the sentence language presentation was blocked. Consequently, participants knew, during the Codeswitched block, that all sentences would include a codeswitch at a certain point. One could argue that switch costs may appear if the sentences are presented in a mixed design. This is indeed possible and remains to be tested. However, it is also possible that no switch costs will emerge even in a mixed design. The comprehension study by Moreno et al. (2002) included a mixed design. The researchers found that the group of more balanced bilingual codeswitchers displayed less switch costs than the less balanced bilinguals. Therefore, it is possible that both groups of bilinguals in the present study, who were very exposed to codeswitching and who codeswitched often in their daily lives, would not exhibit switch costs, even in a mixed design. A second caveat that should be considered is the fact that all the codeswitched sentences began in Spanish and, at some point,
switched into English. Because they were living in the U.S. and studying at a U.S. institution, English was probably the dominant language of these participants, particularly in the case of the early bilinguals. It is, thus, possible that switch costs may emerge when these participants are exposed to sentences that begin in English and switch into their less dominant language, Spanish. Again, this prediction remains to be examined in future studies.
Chapter 6

Conclusions and future directions

This dissertation has set out to fill a gap in the codeswitching literature, namely to examine the comprehension of intrasentential codeswitches. Reports of natural Spanish-English codeswitched discourse have provided evidence of the prevalence of *estar*+English participle switches over *haber*+English participle switches, despite their superficial similarity. The dissimilar distribution of these two types of codeswitches in both oral and written production corpora was confirmed in a corpus study (Chapter 2) as part of this dissertation. Given that *estar*+English participle switches are more frequent than *haber*+English participle switches in production, this dissertation examined if these production patterns were reflected in Spanish-English bilinguals’ comprehension costs. In other words, these studies investigated whether more frequent *estar*+English participle switches were easier to comprehend than less frequent *haber*+English participle switches. Other questions examined in this dissertation were whether age of second language (L2) acquisition and task played a role in the way these two types of codeswitches were processed during reading. Finally, this dissertation examined the extent to which Spanish-English bilinguals exhibited switch costs during the comprehension of these two types of codeswitches. This set of studies focused specifically on reading comprehension of codeswitches, by using the eye-tracking technique. Although codeswitching is primarily an oral phenomenon, it is increasingly common in written language, not only in literary texts, but also in more informal methods of electronic communication, such as email, text messages, and other forms of social media (e.g., blogs, Facebook, Twitter). Therefore, at present, codeswitching in written language is less unexpected and unfamiliar to bilingual readers than it was in the past.
Several groups of Spanish-English bilinguals completed the eye-tracking study, all of which were proficient in both languages and codeswitched on a regular basis. First, a study was conducted in State College, Pennsylvania, in which two groups of Penn State University students participated. One group consisted of bilinguals who had been born and raised in the United States (U.S.) and had acquired both English and Spanish at an early age. These participants were labeled early bilinguals. The other group comprised bilinguals who had been born in Spanish-speaking countries and had moved to the U.S. during adolescence or adulthood. These bilinguals acquired their L2 (English) at a later age and were, therefore, labeled late bilinguals. A second study was conducted in West Harlem, New York, where two groups of very similar early and late bilinguals, who were students at City College of New York, took part in the experiment. The purpose of including these additional two groups of participants was to see whether the same results emerged with bilinguals who belonged to a much larger, established community of bilingual codeswitchers.

The results were very similar for all four groups of bilinguals. In the comprehension task (when participants were asked to read sentences and to answer comprehension questions), they always processed estar+English participle switches more easily than haber+English participle switches. This difference emerged when the mean reading times of the participle in each of these two types of codeswitches were compared to the mean reading times of the participle in baseline codeswitched sentences, in which the switch occurred at the auxiliary. Figures 6-1 and 6-2 present the mean reading times at the participle by condition for the State College participants and the New York participants, respectively.
Figure 6-1: Mean reading times at the participle by condition for the State College (SC) participants during the comprehension task (*p ≤ .05)

Figure 6-2: Mean reading times at the participle by condition for the New York (NY) participants during the comprehension task (*p ≤ .05)
As shown in the figures, the four groups displayed easier processing of the *estar*-English participle switches compared to the *haber*-English participle switches in one or both of the extracted measures of early processing (i.e., gaze duration and regression path time) as well as in the measure of later processing (i.e., total time).

It may seem surprising that no large differences were found between the bilingual groups in State College and those in New York. Not belonging to an established community of bilingual codeswitchers could have meant that the State College bilinguals were less exposed to the unequal distribution of *estar*-English participle switches and *haber*-English participle switches, which is typically found in natural production. However, they proved to be just as sensitive to the different frequency of these two types of codeswitches as the New York bilinguals. It is possible that, although the State College bilinguals do not belong to a large community of codeswitchers, they have developed their own micro-community of codeswitchers in which they are surrounded by Hispanic friends with whom they frequently codeswitch, exposing to the same types of codeswitches as the New York bilinguals.

Another result that may seem unexpected is the lack of large differences between the early bilinguals and the late bilinguals. One prediction was that the late bilinguals would be less sensitive to the different distribution of *estar*-English participle switches and *haber*-English participle switches because they have been exposed to codeswitching and they have been codeswitching themselves for less time than the early bilinguals. Nonetheless, the late bilinguals demonstrated as much sensitivity to the uneven frequency of these two types of codeswitches as the early bilinguals. This was the case both with the State College participants and the New York participants. It is possible that, although the late bilinguals were exposed to codeswitching and began codeswitching themselves later in life, they may have been intensely exposed to codeswitching from the moment that they arrived in the U.S., giving them enough time to
experience the distinct occurrence of these two types of codeswitches in natural production by the
time they participated in the study.

Considering that these four groups of bilinguals were sufficiently exposed to natural
Spanish-English codeswitching production patterns, these results lend support for constraint
satisfaction models of language processing (MacDonald & Seidenberg, 2006), such as the
Production-Distribution-Comprehension account (PDC, Gennari & MacDonald, 2009;
MacDonald & Thornton, 2009). According to the PDC, previous experience with language plays
a crucial role when it comes to language comprehension. Comprehension processes weigh the
probability of alternative interpretations, which are informed by different sources of information
and largely based on prior experience. Therefore, frequency of exposure to different linguistic
structures strongly affects the way those structures are processed when they are encountered in
comprehension.

Although this model has only been used to explain comprehension costs associated with
unilingual structures, it can also account for comprehension of codeswitched structures.
Regarding the codeswitched structures examined here, when describing progressive actions,
estar+English participle codeswitches are produced more frequently than are haber+English
participle codeswitches, when describing perfect actions. This uneven frequency of production
creates distinct distributional patterns for each of these codeswitched structures. These
distributional patterns are learned over time by individuals who are exposed to this input. The
distributional patterns then become the probabilistic constraints that guide comprehension. These
Spanish-English bilinguals have been more frequently exposed to estar+English participle
switches; accordingly, they experience facilitated processing when they face these codeswitches
in reading comprehension. However, they have been rarely, if ever, exposed to haber+English
participle switches, and this is reflected in their difficult processing of these switches during
comprehension.
Another finding that emerged from this dissertation was the lack of switch costs that the two groups of New York bilinguals displayed during the comprehension task. In this case, the comprehension of unilingual and codeswitched sentences was compared. Reading times were measured on the specific word in the sentence where the codeswitch occurred (i.e., the auxiliary). Those reading times were compared to the reading times of the same word in the unilingual sentences. The mean reading times at the auxiliary by condition for the New York participants are exhibited in Figure 6-3.

![Figure 6-3: Mean reading times at the auxiliary by condition for the New York (NY) participants during the comprehension task](image)

When reading for comprehension, both early and late bilinguals read the auxiliary in unilingual sentences and the auxiliary in codeswitched sentences at a similar speed. The general finding among studies that have examined switch costs during production and comprehension is that switch costs consistently appear in bilingual processing. The results in this dissertation seem to
contradict those found in other switch cost studies. However, one has to consider that the type of codeswitching examined in most of the earlier studies is very different from the type of codeswitching examined here. Most of the studies that focus on production (e.g., Meuter & Allport, 1999) use the cued switching procedure, in which participants produce isolated lexical items and must switch languages when they are cued by a specific visual or auditory signal. This type of codeswitching differs greatly from more natural codeswitching, in which switch costs are not found (Timm, 1983). Moreover, the studies that focus on comprehension (e.g., Moreno, Federmeier, & Kutas, 2002) use sentences with single-word switches. Again, that type of codeswitches is different from the intrasentential codeswitches examined in this dissertation. In fact, studies (e.g., Poplack, 1980) have shown that more balanced bilinguals, such as those who took part in the present study, tend to produce intrasentential switches more often than single-word insertions. These differences between earlier switch cost studies and the present study may explain the contradictory results. The types of codeswitches that the participants encountered during reading (i.e., switches at a phrasal boundary) are probably switches that they recurrently use and to which they are frequently exposed, and, therefore, do not result in any additional costs compared to their unilingual counterparts.

The results from the studies in this dissertation also yielded clear task effects. The acceptability judgment task (the task in which participants were asked to read the codeswitched sentences and to judge if they were acceptable or not) brought about processing differences between the early bilinguals and the late bilinguals. When reading to perform acceptability judgments, the two groups of early bilinguals no longer displayed sensitivity to the different distribution of the estar+English participle switches and the haber+English participle switches, which is exhibited in natural production. Instead, they displayed more difficulty processing both types of switches at the participle, when each of these was compared to its corresponding switch at the auxiliary. These results are displayed by the mean reading times at the participle for each
experimental condition, as shown in Figure 6-4. Therefore, here the particular type of auxiliary (i.e., progressive or perfect) involved in the switch did not matter; all switches at the participle were processed with increased difficulty.

The late bilinguals’ processing was not significantly affected by the task. The New York late bilinguals displayed the same processing patterns in the comprehension task and the acceptability judgment task. They consistently processed the haber+English participle switches with more difficulty than the estar+English participle switches, as displayed in measures of early and later processing in Figure 6-5.

Figure 6-4: Mean reading times at the participle by condition for the early bilinguals in State College (SC) and New York (NY) during the acceptability judgment task (* p ≤ .05)
The only difference that the State College late bilinguals exhibited between the two tasks was the particular moment at which they displayed distinct processing of *estar*+English participle switches and *haber*+English participles switches. During the comprehension task, they exhibited different processing of these two types of codeswitches in a measure of early processing (i.e., regression path) and a measure of later processing (i.e., total time). During the acceptability judgment task, they no longer processed these two codeswitches differently in any measure of early processing, but only in the measure of later processing. Although it took longer, ultimately the State College late bilinguals still processed the *estar*+English participle switches and the *haber*+English participle switches differently during the acceptability judgment task. Therefore, while the way in which the early bilinguals processed these codeswitches was largely influenced by the task that they were given, the late bilinguals’ processing seemed to be unaffected by the task. Although the early and late bilinguals differed in this respect, when we consider the overt acceptability judgments, the groups resembled each other again.

Figure 6-5: Mean reading times at the participle by condition for the late bilinguals in State College (SC) and New York (NY) during the acceptability judgment task (* p ≤ .05)
judgments were not identical for all groups of participants, but they did follow the same pattern. Switches at the auxiliary and *estarn*+English participle switches received higher proportions of acceptable judgments while *habern*+English participle switches received the lowest proportion of acceptable judgments.

The reason why the early bilinguals and the late bilinguals were affected differently by the task still remains unexplained. Although only speculative, a possible explanation may have to do with their perception of codeswitching. The early bilinguals may have been raised in family circles in which codeswitching, although present, was frowned upon and maybe even corrected. That is, the early bilinguals may associate codeswitching with an incorrect way of speaking, particularly when completing a task that relies on metalinguistic judgments. Their processing, then, may reflect the idea that codeswitching is incorrect, especially more complex types of codeswitches, such as those that occur within phrasal boundaries (i.e., switches at the participle). The late bilinguals, however, may have not had that experience and, therefore, may be, in a way, less aware of the stigma that is sometimes attached to codeswitching. Consequently, their reading patterns may be unaffected by the task, resulting in equal processing during the comprehension task and the acceptability judgment task.

The stigma that is occasionally still linked to codeswitching is a remnant of the idea that codeswitching signals lack of proficiency in both languages and is randomly recurred to during speech, whenever bilinguals are unable to express themselves fully in one language. However, the results in this dissertation show that codeswitching is not haphazard, but instead very systematic. Bilinguals do not switch anywhere in the sentence. The points at which they choose to codeswitch are quite controlled. It is because of this systematicity in production that the distributional patterns of the different types of codeswitches emerge and are subsequently reflected in the way the different types of codeswitches are processed. If codeswitching were random, there would not be differences in the distribution of codeswitches, and all types of
codeswitches would be processed in the same way. The fact that the bilinguals in these studies processed the more frequent types of codeswitches with more ease than the less frequent types of codeswitches (at least in the comprehension task) shows that these systematic differences in production do exist and that the distributional patterns that are created as a result are used during comprehension to facilitate processing. In that sense, codeswitching seems to function just as systematically as unilingual discourse does.

This dissertation should not be viewed as a final product, but rather as a work in progress. As such, it has several limitations that can be addressed in future studies. For instance, the studies in this dissertation have examined reading comprehension of codeswitches. However, codeswitching, although it is increasingly common in written language, is still more frequently used in speech than in writing. It is, therefore, important to test the comprehension of estar+English participle switches and haber+English participle switches in a more ecologically valid scenario in which the codeswitches are presented to the participants auditorily. Being exposed to oral codeswitches may seem more natural to bilingual codeswitchers. The results of an auditory comprehension study should also display, maybe even more clearly, easier processing of the more frequent estar+English participle switches than the less common haber+English participle switches.

Another future avenue for this research is related to switch costs. In this dissertation, all the codeswitched sentences were presented to participants in a single block. Therefore, participants expected to encounter a codeswitch as they read every sentence. They did not know exactly where in each sentence the codeswitch would appear, but they knew that there would always be one. It is possible that the blocked-presentation design was responsible for the absence of switch costs exhibited by the participants during the comprehension task. An important follow-up study would be to examine the comprehension of these two types of codeswitches in a mixed-presentation design, in which unilingual and codeswitched sentences are pseudo-randomly
interleaved. This design would allow us to examine whether bilingual codeswitchers exhibit
switch costs in an environment in which the codeswitches are more unexpected. This type of
presentation would also better resemble natural codeswitched discourse, in which codeswitching
does not occur in every single sentence, but is instead more dispersed throughout the discourse.

Finally, an additional way in which this dissertation work can be expanded is related to
the inclusion of an additional group of bilingual participants. This set of studies has shown that
there was a direct relation between the bilingual codeswitchers’ previous experience with the
distinct distributional patterns of estar+English participles and haber+English participles and the
way that they processed these two types of switches during comprehension. In other words, these
bilinguals had been exposed to the estar+English participle switches more often than they had
been exposed to the haber+English participle switches, and this was reflected in their
comprehension costs – they processed the former type of switch more easily than the latter type.
A supplemental study would entail examining how bilinguals who do not codeswitch and who
belong to a community in which they are not exposed to Spanish-English codeswitching patterns
process these two types of codeswitches. If constraint satisfaction models, such as the PDC,
accurately describe comprehension processes, then we can hypothesize that bilinguals who
belong to a community of non-codeswitchers should process these two types of codeswitches
similarly.
References


Oxford, United Kingdom: Claredon Press.


Appendix A

Sample questions from the Language History Questionnaire

(1) Did you begin to speak both English and Spanish before age 5?    Yes    No

(2) In general, which language do you prefer to use?
    English    Spanish    Both    It depends on whom I talk to

(3) How many hours per day do you read in English?
    30 minutes    1 hour    About 2 hours    About 3 hours    More than 3 hours

(4) On a scale from 1 (almost never) to 10 (always), rate to what extent you are currently exposed to Spanish while reading.
    1    2    3    4    5    6    7    8    9    10

(5) What type of materials do you read in Spanish? (choose all that apply)
    Newspaper    Magazines    Textbooks    Literary works    Other: __________

(6) Codeswitching means using both Spanish and English in the same sentence when you are talking to someone else. Do you ever codeswitch?    Yes    No

(7) Why do you think you codeswitch?
Appendix B

Examples of the outline drawings included in the Boston Naming Vocabulary Test

Figure 1
Picture of a whistle

Figure 2
Picture of an octopus

In the case of Figure 1, which belongs to the English language block of the Boston Naming Vocabulary Test (BNT), participants were expected to say “whistle.” In the case of Figure 2, which belongs to the Spanish block, participants were expected to say “pulpo” (‘octopus’).
Appendix C

Sample items from the grammatical competency tests

The Michigan English Language Institute College Entrance Test (MELICET) that the participants completed was divided in two parts. In the first part, participants were presented with different cases of fictitious conversations and were instructed to choose the word or phrase that best completed the conversation. An example of this exercise is provided below (the correct answer is marked with an asterisk).

“What is that thing?”

“That _____ a spider.”

a) to call          b) for calling          c) be called          *d) is called

In the second part of the MELICET, a text was presented in which several sentences contained a blank. Participants were asked to read the passage and to select the word that best fit the blank in both grammar and meaning, as displayed below.

Long ago, roads were only trails for people and animals to walk on, but today roads must be made for cars, trucks, and busses. The most modern type of _____, the highway, is a marvel of modern engineering.

a) way          *b) road          c) travel          d) superhighway
The Diplomas de Español como Lengua Extranjera (Diplomas of Spanish as a Foreign Language, DELE) test that the participants completed was divided into three parts. The first part consisted of a text that included several sentences with blanks. Participants were asked to fill the blanks with one of the provided options. An example of this exercise is displayed below (the correct answer is marked with an asterisk).

Los padres siempre se están preguntando cómo conseguir que sus hijos sean unos niños talentosos y sanos y las soluciones pueden estar más cerca de lo que creemos. Ni tónicos, ni vitaminas, ni cursos de lectura veloz pueden conseguir tantos resultados en los niños _____ la práctica constante de hábitos saludables.

a) que *b) como c) cuales

‘Parents are always asking themselves how to get their children to be talented and healthy and the answers may be closer than we think. Nor tonics, nor vitamins nor speed reading courses can provide as many results in children _____ the constant practice of healthy habits.

a) that *b) as c) which’
The second part of the DELE contained individual sentences that included a bolded word or phrase. Participants were instructed to select the best definition for the bolded words from the provided options.

Tengo la impresión de que los libros que yo tenía de pequeña están **dispersos** por la casa de mis padres.

a) ordenados  b) perdidos  *c) esparcidos

‘(I) have the impression that the books that I had as a child are **dispersed** through my parents’ house.

a) ordered  b) lost  *c) scattered’

In the third part of the DELE, participants were presented with individual sentences that contained a blank and they were instructed to select the word that best fit the blank in both grammar and meaning, as displayed below.

En la compañía se está decidiendo estos días si _____ nuevos horarios para los trabajadores.

a) haya    *b) habrá

‘These days the company is deciding whether _____ new schedules for the workers.

a) there were    *b) there will be’
Appendix D

Codeswitched experimental item sets

El abogado cree que sus clientes are notifying the police of the events.
El abogado cree que sus clientes están notifying the police of the events.
El abogado cree que sus clientes have notified the police of the events.
El abogado cree que sus clientes han notified the police of the events.

El abogado descubrió que los criminales are bribing the policeman to destroy the evidence.
El abogado descubrió que los criminales están bribing the policeman to destroy the evidence.
El abogado descubrió que los criminales have bribed the policeman to destroy the evidence.
El abogado descubrió que los criminales han bribed the policeman to destroy the evidence.

El abogado garantizó que los criminales are improving their behavior in jail.
El abogado garantizó que los criminales están improving their behavior in jail.
El abogado garantizó que los criminales have improved their behavior in jail.
El abogado garantizó que los criminales han improved their behavior in jail.

El actor notó que los asistentes are carrying the suitcases to his hotel room.
El actor notó que los asistentes están carrying the suitcases to his hotel room.
El actor notó que los asistentes have carried the suitcases to his hotel room.
El actor notó que los asistentes han carried the suitcases to his hotel room.

El alcalde confirmó que los detectives are negotiating the release of the hostages.
El alcalde confirmó que los detectives están negotiating the release of the hostages.
El alcalde confirmó que los detectives have negotiated the release of the hostages.
El alcalde confirmó que los detectives han negotiated the release of the hostages.

El arquitecto supone que los pintores are considering the colors for the house.
El arquitecto supone que los pintores están considering the colors for the house.
El arquitecto supone que los pintores have considered the colors for the house.
El arquitecto supone que los pintores han considered the colors for the house.
El carcelero dijo que los prisioneros están lavando su ropa para la semana.
El carcelero dijo que los prisioneros ha lavado su ropa para la semana.

El chef piensa que los turistas están disfrutando de la comida en su restaurante gourmet.
El chef piensa que los turistas han disfrutado de la comida en su restaurante gourmet.

El compositor dice que los pianistas están practicando la sinfonía para la concierto.
El compositor dice que los pianistas han practicado la sinfonía para la concierto.

El comprador piensa que los artistas están pintando los retratos para su colección.
El comprador piensa que los artistas han pintado los retratos para su colección.

El consejero dijo que sus estudiantes están presentando sus resultados en la conferencia.
El consejero dijo que sus estudiantes han presentado sus resultados en la conferencia.

El conserje notó que los instructores están limpiando las mesas en sus salas de clases.
El conserje notó que los instructores han limpiado las mesas en sus salas de clases.

El contador cree que los banqueros están negociando el préstamo para los clientes.
El contador cree que los banqueros han negociado el préstamo para los clientes.

El decano notó que las recepcionistas están filando las solicitudes en orden alfabético.
El decano notó que las recepcionistas han filmado las solicitudes en orden alfabético.
El director afirmó que los técnicos están reparando los fotocopiadores en la biblioteca escolar.
El director afirmó que los técnicos han reparado los fotocopiadores en la biblioteca escolar.

El director confirmó que los actores están ensayando sus líneas para la película.
El director confirmó que los actores han ensayado sus líneas para la película.

El director cree que los editores están modificando el título del libro.
El director cree que los editores han modificado el título del libro.

El director dijo que los instructores están preparando el examen para los estudiantes.
El director dijo que los instructores han preparado el examen para los estudiantes.

El director mencionó que los instructores están reclutando a algunos estudiantes para el club español.
El director mencionó que los instructores han reclutado a algunos estudiantes para el club español.

El dueño dijo que los arquitectos están firmando los documentos para la construcción.
El dueño dijo que los arquitectos han firmado los documentos para la construcción.

El editor notó que los voluntarios están arreglando las fotografías para la sección de Entretenimiento.
El editor notó que los voluntarios han arreglado las fotografías para la sección de Entretenimiento.

El ejecutivo piensa que las recepcionistas están grapando los documentos para su presentación.
El ejecutivo piensa que las recepcionistas han grapado los documentos para su presentación.

El empleado supone que sus colegas están informando a su jefe del accidente.
El empleado supone que sus colegas han informado a su jefe del accidente.
El entrenador dijo que los atletas están ignoring the remarks from the opposing team.
El entrenador dijo que los atletas están ignoring the remarks from the opposing team.
El entrenador dijo que los atletas have ignored the remarks from the opposing team.
El entrenador dijo que los atletas han ignored the remarks from the opposing team.

El entrenador indicó que los gimnastas están performing their routines successfully.
El entrenador indicó que los gimnastas están performing their routines successfully.
El entrenador indicó que los gimnastas have performed their routines successfully.
El entrenador indicó que los gimnastas han performed their routines successfully.

El entrenador mencionó que los atletas están practicing five hours a day.
El entrenador mencionó que los atletas están practicing five hours a day.
El entrenador mencionó que los atletas have practiced five hours a day.
El entrenador mencionó que los atletas han practiced five hours a day.

El entrenador notó que los atletas están grabbing their uniforms from the pile.
El entrenador notó que los atletas están grabbing their uniforms from the pile.
El entrenador notó que los atletas have grabbed their uniforms from the pile.
El entrenador notó que los atletas han grabbed their uniforms from the pile.

El entrenador piensa que los atletas are celebrating their win at the bar.
El entrenador piensa que los atletas están celebrating their win at the bar.
El entrenador piensa que los atletas have celebrated their win at the bar.
El entrenador piensa que los atletas han celebrated their win at the bar.

El estudiante notó que los biólogos están classifying the plants into ten groups.
El estudiante notó que los biólogos están classifying the plants into ten groups.
El estudiante notó que los biólogos have classified the plants into ten groups.
El estudiante notó que los biólogos han classified the plants into ten groups.

El general mencionó que los veteranos están enjoying the celebration in their honor.
El general mencionó que los veteranos están enjoying the celebration in their honor.
El general mencionó que los veteranos have enjoyed the celebration in their honor.
El general mencionó que los veteranos han enjoyed the celebration in their honor.

El gerente notó que los turistas están enjoying their stay in the hotel.
El gerente notó que los turistas están enjoying their stay in the hotel.
El gerente notó que los turistas have enjoyed their stay in the hotel.
El gerente notó que los turistas han enjoyed their stay in the hotel.

El guardia dijo que los prisioneros están cooking the food in the kitchen.
El guardia dijo que los prisioneros están cooking the food in the kitchen.
El guardia dijo que los prisioneros have cooked the food in the kitchen.
El guardia dijo que los prisioneros han cooked the food in the kitchen.

El ingeniero supone que los arquitectos están approving the plans for the project.
El ingeniero supone que los arquitectos están approving the plans for the project.
El ingeniero supone que los arquitectos have approved the plans for the project.
El ingeniero supone que los arquitectos han approved the plans for the project.

El investigador piensa que los gángsters are shipping the drugs to New York City.
El investigador piensa que los gángsters están shipping the drugs to New York City.
El investigador piensa que los gángsters have shipped the drugs to New York City.
El investigador piensa que los gángsters han shipped the drugs to New York City.
El jefe anunció que las secretarias are notifying the media about the strike.
El jefe anunció que las secretarias están notifying the media about the strike.
El jefe anunció que las secretarias have notified the media about the strike.
El jefe anunció que las secretarias han notified the media about the strike.

El jefe dijo que los asistentes are arranging the flowers for the party.
El jefe dijo que los asistentes están arranging the flowers for the party.
El jefe dijo que los asistentes have arranged the flowers for the party.
El jefe dijo que los asistentes han arranged the flowers for the party.

El jefe indicó que los fotógrafos are selecting the pictures for the advertisement.
El jefe indicó que los fotógrafos están selecting the pictures for the advertisement.
El jefe indicó que los fotógrafos have selected the pictures for the advertisement.
El jefe indicó que los fotógrafos han selected the pictures for the advertisement.

El jefe piensa que las recepcionistas are calling the applicants for interviews.
El jefe piensa que las recepcionistas están calling the applicants for interviews.
El jefe piensa que las recepcionistas have called the applicants for interviews.
El jefe piensa que las recepcionistas han called the applicants for interviews.

El juez anunció que los criminales are confessing their involvement in the kidnapping.
El juez anunció que los criminales están confessing their involvement in the kidnapping.
El juez anunció que los criminales have confessed their involvement in the kidnapping.
El juez anunció que los criminales han confessed their involvement in the kidnapping.

El locutor dijo que los entrenadores are reaching their goals with the players.
El locutor dijo que los entrenadores están reaching their goals with the players.
El locutor dijo que los entrenadores have reached their goals with the players.
El locutor dijo que los entrenadores han reached their goals with the players.

El maestro asumió que los estudiantes are memorizing the information in the first chapter.
El maestro asumió que los estudiantes están memorizing the information in the first chapter.
El maestro asumió que los estudiantes have memorized the information in the first chapter.
El maestro asumió que los estudiantes han memorized the information in the first chapter.

El maestro notó que los estudiantes are copying the answers on their desks.
El maestro notó que los estudiantes están copying the answers on their desks.
El maestro notó que los estudiantes have copied the answers on their desks.
El maestro notó que los estudiantes han copied the answers on their desks.

El organizador confirmó que los voluntarios are cleaning the ballroom for the reception.
El organizador confirmó que los voluntarios están cleaning the ballroom for the reception.
El organizador confirmó que los voluntarios have cleaned the ballroom for the reception.
El organizador confirmó que los voluntarios han cleaned the ballroom for the reception.

El periodista anunció que los músicos are producing the album in the studio.
El periodista anunció que los músicos están producing the album in the studio.
El periodista anunció que los músicos have produced the album in the studio.
El periodista anunció que los músicos han produced the album in the studio.
El portavoz dijo que los chefs are cooking the meal for the presidential visit.
El portavoz dijo que los chefs están cooking the meal for the presidential visit.
El portavoz dijo que los chefs ha
cooked the meal for the presidential visit.
El portavoz dijo que los chefs han cooked the meal for the presidential visit.

El presidente anunció que los senadores are negotiating the terms of the agreement.
El presidente anunció que los senadores están negotiating the terms of the agreement.
El presidente anunció que los senadores have negotiated the terms of the agreement.
El presidente anunció que los senadores han negotiated the terms of the agreement.

El presidente aseguró que los soldados are protecting the citizens from the attacks.
El presidente aseguró que los soldados están protecting the citizens from the attacks.
El presidente aseguró que los soldados have protected the citizens from the attacks.
El presidente aseguró que los soldados han protected the citizens from the attacks.

El psiquiatra afirmó que los prisioneros are justifying their behavior in the session.
El psiquiatra afirmó que los prisioneros están justifying their behavior in the session.
El psiquiatra afirmó que los prisioneros have justified their behavior in the session.
El psiquiatra afirmó que los prisioneros han justified their behavior in the session.

El reportero afirmó que los senadores are authorizing the taxes on gas.
El reportero afirmó que los senadores están authorizing the taxes on gas.
El reportero afirmó que los senadores have authorized the taxes on gas.
El reportero afirmó que los senadores han authorized the taxes on gas.

El reportero confirmó que los senadores are requesting the funds for the project.
El reportero confirmó que los senadores están requesting the funds for the project.
El reportero confirmó que los senadores have requested the funds for the project.
El reportero confirmó que los senadores han requested the funds for the project.

El reportero descubrió que los senadores are approving the changes to the bill.
El reportero descubrió que los senadores están approving the changes to the bill.
El reportero descubrió que los senadores have approved the changes to the bill.
El reportero descubrió que los senadores han approved the changes to the bill.

El reportero dijo que las modelos are signing a contract with the agency.
El reportero dijo que las modelos están signing a contract with the agency.
El reportero dijo que las modelos have signed a contract with the agency.
El reportero dijo que las modelos han signed a contract with the agency.

El reportero dijo que los activistas are requesting a meeting with the president.
El reportero dijo que los activistas están requesting a meeting with the president.
El reportero dijo que los activistas have requested a meeting with the president.
El reportero dijo que los activistas han requested a meeting with the president.

El sargento garantizó que los detectives are removing the evidence from the crime scene.
El sargento garantizó que los detectives están removing the evidence from the crime scene.
El sargento garantizó que los detectives have removed the evidence from the crime scene.
El sargento garantizó que los detectives han removed the evidence from the crime scene.

El sargento garantizó que los soldados are preparing the weapons for the mission.
El sargento garantizó que los soldados están preparing the weapons for the mission.
El sargento garantizó que los soldados have prepared the weapons for the mission.
El sargento garantizó que los soldados han prepared the weapons for the mission.
El supervisor confirmó que los jardineros están watering the plants in the yard.
El supervisor confirmó que los jardineros have watered the plants in the yard.
El supervisor confirmó que los jardineros han watered the plants in the yard.

El supervisor dijo que los carpinteros están finishing the balcony of the house.
El supervisor dijo que los carpinteros have finished the balcony of the house.
El supervisor dijo que los carpinteros han finished the balcony of the house.

El supervisor garantizó que los empleados están fixing the mistakes in the reports.
El supervisor garantizó que los empleados have fixed the mistakes in the reports.
El supervisor garantizó que los empleados han fixed the mistakes in the reports.

El supervisor mencionó que los carpinteros están fixing the cabinets in the kitchen.
El supervisor mencionó que los carpinteros have fixed the cabinets in the kitchen.
El supervisor mencionó que los carpinteros han fixed the cabinets in the kitchen.

El supervisor piensa que los empleados están discussing the issues in their meeting.
El supervisor piensa que los empleados have discussed the issues in their meeting.
El supervisor piensa que los empleados han discussed the issues in their meeting.

El técnico garantizó que los empleados están removing the viruses from the computers.
El técnico garantizó que los empleados have removed the viruses from the computers.
El técnico garantizó que los empleados han removed the viruses from the computers.

El vendedor confirmó que los coleccionistas están importing the sculptures from India.
El vendedor confirmó que los coleccionistas have imported the sculptures from India.
El vendedor confirmó que los coleccionistas han imported the sculptures from India.

La compañía supone que los supervisores están authorizing the purchase of new equipment.
La compañía supone que los supervisores have authorized the purchase of new equipment.
La compañía supone que los supervisores han authorized the purchase of new equipment.

La enfermera afirmó que los doctores están consulting a specialist about the results.
La enfermera afirmó que los doctores have consulted a specialist about the results.
La enfermera afirmó que los doctores han consulted a specialist about the results.

La enfermera confirmó que los doctores están importing the medicine from China.
La enfermera confirmó que los doctores have imported the medicine from China.
La enfermera confirmó que los doctores han imported the medicine from China.
La enfermera descubrió que los cirujanos están deceiving the patient about his illness.
La enfermera descubrió que los cirujanos ha deceived the patient about his illness.
La enfermera piensa que los cirujanos are approving the procedure for the patient.
La enfermera piensa que los cirujanos ha approved the procedure for the patient.
La estilista confirmó que los diseñadores are organizing their collections for the fashion show.
La estilista confirmó que los diseñadores ha organized their collections for the fashion show.
La familia notó que los jardineros are planting the trees in the backyard.
La familia notó que los jardineros ha planted the trees in the backyard.
La maestra piensa que los tutores are helping the students with their homework.
La maestra piensa que los tutores ha helped the students with their homework.
La monja piensa que los turistas are dirtying the cathedral during their visit.
La monja piensa que los turistas ha dirtied the cathedral during their visit.
La novia cree que los coordinadores are arranging the seating for the wedding reception.
La novia cree que los coordinadores ha arranged the seating for the wedding reception.
La prensa confirmó que los actores are donating the money to charity.
La prensa confirmó que los actores ha donated the money to charity.
La profesora anunció que los editores están approving her article for the journal.
La profesora anunció que los editores están approving her article for the journal.
La profesora anunció que los editores have approved her article for the journal.
La profesora anunció que los editores han approved her article for the journal.

La radio anunció que los guitarristas están organizing a concert in the convention center.
La radio anunció que los guitarristas están organizing a concert in the convention center.
La radio anunció que los guitarristas have organized a concert in the convention center.
La radio anunció que los guitarristas han organized a concert in the convention center.

La reportera afirmó que los científicos están testing the vaccine on rats.
La reportera afirmó que los científicos están testing the vaccine on rats.
La reportera afirmó que los científicos have tested the vaccine on rats.
La reportera afirmó que los científicos han tested the vaccine on rats.

La revista indicó que los actores están answering the letters from their fans.
La revista indicó que los actores están answering the letters from their fans.
La revista indicó que los actores have answered the letters from their fans.
La revista indicó que los actores han answered the letters from their fans.

La secretaria dijo que los asistentes están accusing their boss of fraud.
La secretaria dijo que los asistentes están accusing their boss of fraud.
La secretaria dijo que los asistentes have accused their boss of fraud.
La secretaria dijo que los asistentes han accused their boss of fraud.

La secretaria notó que los asistentes están removing the trash from the office.
La secretaria notó que los asistentes están removing the trash from the office.
La secretaria notó que los asistentes have removed the trash from the office.
La secretaria notó que los asistentes han removed the trash from the office.

La secretaria piensa que los asistentes están dividing the files into three piles.
La secretaria piensa que los asistentes están dividing the files into three piles.
La secretaria piensa que los asistentes have divided the files into three piles.
La secretaria piensa que los asistentes han divided the files into three piles.

La superintendente garantiza que los instructores están testing the students appropriately.
La superintendente garantiza que los instructores están testing the students appropriately.
La superintendente garantiza que los instructores have tested the students appropriately.
La superintendente garantiza que los instructores han tested the students appropriately.

Los diseñadores notaron que las modelos están practicing their walk on the runway.
Los diseñadores notaron que las modelos están practicing their walk on the runway.
Los diseñadores notaron que las modelos have practiced their walk on the runway.
Los diseñadores notaron que las modelos han practiced their walk on the runway.

Los estudiantes notaron que las profesoras están placing their quizzes on the bookshelf.
Los estudiantes notaron que las profesoras están placing their quizzes on the bookshelf.
Los estudiantes notaron que las profesoras have placed their quizzes on the bookshelf.
Los estudiantes notaron que las profesoras han placed their quizzes on the bookshelf.
Los inquilinos notaron que los electricistas están fixing the powerlines in the building.
Los inquilinos notaron que los electricistas han fixed the powerlines in the building.

Los niños piensan que los magos are turning the rabbits into doves.
Los niños piensan que los magos han turned the rabbits into doves.

Mi vecino notó que los estudiantes are watching the game on TV.
Mi vecino notó que los estudiantes han watched the game on TV.
Appendix E

Unilingual experimental item sets

The lawyer believes that his clients are notifying the police of the events.
The lawyer believes that his clients have notified the police of the events.

The lawyer discovered that the criminals are bribing the policeman to destroy the evidence.
The lawyer discovered that the criminals have bribed the policeman to destroy the evidence.

The lawyer guaranteed that the criminals are improving their behavior in jail.
The lawyer guaranteed that the criminals have improved their behavior in jail.

The actor noticed that the assistants are carrying the suitcases to his hotel room.
The actor noticed that the assistants have carried the suitcases to his hotel room.

The agent assumes that the tourists are returning the car at the airport.
The agent assumes that the tourists have returned the car at the airport.

The mayor confirmed that the detectives are negotiating the release of the hostages.
The mayor confirmed that the detectives have negotiated the release of the hostages.

The mayor said that the sculptors are creating a statue for the park.
The mayor said that the sculptors have created a statue for the park.

The mayor noticed that the painters are changing the design of the mural.
The mayor noticed that the painters have changed the design of the mural.

The architect supposes that the painters are considering the colors for the house.
The architect supposes that the painters have considered the colors for the house.

The guard said that the prisoners are washing their clothes for the week.
The guard said that the prisoners have washed their clothes for the week.

The chef thinks that the tourists are enjoying the food at his gourmet restaurant.
The chef thinks that the tourists have enjoyed the food at his gourmet restaurant.

The composer says that the pianists are practicing the symphony for the concert.
The composer says that the pianists have practiced the symphony for the concert.

The composer said that the pianists are modifying the music for the concert.
The composer said that the pianists have modified the music for the concert.

The composer supposes that the musicians are tuning their instruments before the concert.
The composer supposes that the musicians have tuned their instruments before the concert.

The buyer thinks that the artists are painting the portraits for his collection.
The buyer thinks that the artists have painted the portraits for his collection.

The advisor said that his students are presenting their results at the conference.
The advisor said that his students have presented their results at the conference.
The janitor noticed that the instructors are wiping the desks in their classrooms.  
The janitor noticed that the instructors have wiped the desks in their classrooms.

The accountant believes that the bankers are negotiating the loan for the clients.  
The accountant believes that the bankers have negotiated the loan for the clients.

The accountant thinks that the bankers are preparing the report for the supervisors.  
The accountant thinks that the bankers have prepared the report for the supervisors.

The dean noticed that the receptionists are filing the applications in alphabetical order.  
The dean noticed that the receptionists have filed the applications in alphabetical order.

The principal affirmed that the technicians are repairing the photocopiers in the school library.  
The principal affirmed that the technicians have repaired the photocopiers in the school library.

The director confirmed that the actors are rehearsing their lines for the movie.  
The director confirmed that the actors have rehearsed their lines for the movie.

The director believes that the editors are modifying the title of the book.  
The director believes that the editors have modified the title of the book.

The principal said that the instructors are preparing the exam for the students.  
The principal said that the instructors have prepared the exam for the students.

The director said that the producers are preparing the set for the movie.  
The director said that the producers have prepared the set for the movie.

The director mentioned that the instructors are recruiting some students for the Spanish club.  
The director mentioned that the instructors have recruited some students for the Spanish club.

The owner said that the architects are signing the documents for the construction.  
The owner said that the architects have signed the documents for the construction.

The editor noticed that the volunteers are arranging the photographs for the Entertainment section.  
The editor noticed that the volunteers have arranged the photographs for the Entertainment section.

The executive thinks that the receptionists are stapling the documents for his presentation.  
The executive thinks that the receptionists have stapled the documents for his presentation.

The employee supposes that his colleagues are notifying their boss of the accident.  
The employee supposes that his colleagues have notified their boss of the accident.

The coach said that the athletes are ignoring the remarks from the opposing team.  
The coach said that the athletes have ignored the remarks from the opposing team.

The coach indicated that the gymnasts are performing their routines successfully.  
The coach indicated that the gymnasts have performed their routines successfully.

The coach mentioned that the athletes are practicing five hours a day.  
The coach mentioned that the athletes have practiced five hours a day.

The coach noticed that the athletes are grabbing their uniforms from the pile.  
The coach noticed that the athletes have grabbed their uniforms from the pile.
The coach thinks that the athletes are celebrating their win at the bar.
The coach thinks that the athletes have celebrated their win at the bar.

The student noticed that the biologists are classifying the plants into ten groups.
The student noticed that the biologists have classified the plants into ten groups.

The general mentioned that the veterans are enjoying the celebration in their honor.
The general mentioned that the veterans have enjoyed the celebration in their honor.

The manager noticed that the tourists are enjoying their stay in the hotel.
The manager noticed that the tourists have enjoyed their stay in the hotel.

The guard said that the prisoners are cooking the food in the kitchen.
The guard said that the prisoners have cooked the food in the kitchen.

The engineer supposes that the architects are approving the plans for the project.
The engineer supposes that the architects have approved the plans for the project.

The investigator thinks that the gangsters are shipping the drugs to New York City.
The investigator thinks that the gangsters have shipped the drugs to New York City.

The boss announced that the secretaries are notifying the media about the strike.
The boss announced that the secretaries have notified the media about the strike.

The boss said that the assistants are arranging the flowers for the party.
The boss said that the assistants have arranged the flowers for the party.

The judge announced that the criminals are confessing their involvement in the kidnapping.
The judge announced that the criminals have confessed their involvement in the kidnapping.

The announcer said that the trainers are reaching their goals with the players.
The announcer said that the trainers have reached their goals with the players.

The teacher assumed that the students are memorizing the information in the first chapter.
The teacher assumed that the students have memorized the information in the first chapter.

The doctor said that the patients are recovering from the operation.
The doctor said that the patients have recovered from the operation.

The bookkeeper noticed that the clerks are filing the documents on the shelves.
The bookkeeper noticed that the clerks have filed the documents on the shelves.
The president announced that the senators are negotiating the terms of the agreement. The president announced that the senators have negotiated the terms of the agreement.

The president assured that the soldiers are protecting the citizens from the attacks. The president assured that the soldiers have protected the citizens from the attacks.

The psychiatrist affirmed that the prisoners are justifying their behavior in the session. The psychiatrist affirmed that the prisoners have justified their behavior in the session.

The reporter affirmed that the senators are authorizing the taxes on gas. The reporter affirmed that the senators have authorized the taxes on gas.

The reporter confirmed that the senators are requesting the funds for the project. The reporter confirmed that the senators have requested the funds for the project.

The reporter discovered that the senators are approving the changes to the bill. The reporter discovered that the senators have approved the changes to the bill.

The reporter said that the models are signing a contract with the agency. The reporter said that the models have signed a contract with the agency.

The reporter said that the activists are requesting a meeting with the president. The reporter said that the activists have requested a meeting with the president.

The sergeant guaranteed that the detectives are removing the evidence from the crime scene. The sergeant guaranteed that the detectives have removed the evidence from the crime scene.

The sergeant guaranteed that the soldiers are preparing the weapons for the mission. The sergeant guaranteed that the soldiers have prepared the weapons for the mission.

The supervisor confirmed that the gardeners are watering the plants in the yard. The supervisor confirmed that the gardeners have watered the plants in the yard.

The supervisor said that the carpenters are finishing the balcony of the house. The supervisor said that the carpenters have finished the balcony of the house.

The supervisor guaranteed that the employees are fixing the mistakes in the reports. The supervisor guaranteed that the employees have fixed the mistakes in the reports.

The supervisor mentioned that the carpenters are fixing the cabinets in the kitchen. The supervisor mentioned that the carpenters have fixed the cabinets in the kitchen.

The supervisor thinks that the employees are discussing the issues in their meeting. The supervisor thinks that the employees have discussed the issues in their meeting.

The technician guaranteed that the employees are removing the viruses from the computers. The technician guaranteed that the employees have removed the viruses from the computers.

The seller confirmed that the collectors are importing the sculptures from India. The seller confirmed that the collectors have imported the sculptures from India.
The company supposes that the supervisors are authorizing the purchase of new equipment.  
The company supposes that the supervisors have authorized the purchase of new equipment.

The nurse affirmed that the doctors are consulting a specialist about the results.  
The nurse affirmed that the doctors have consulted a specialist about the results.

The nurse confirmed that the doctors are importing the medicine from China.  
The nurse confirmed that the doctors have imported the medicine from China.

The nurse discovered that the surgeons are deceiving the patient about his illness.  
The nurse discovered that the surgeons have deceived the patient about his illness.

The nurse thinks that the surgeons are approving the procedure for the patient.  
The nurse thinks that the surgeons have approved the procedure for the patient.

The nurse thinks that the surgeons are authorizing the treatment for the patient.  
The nurse thinks that the surgeons have authorized the treatment for the patient.

The stylist confirmed that the designers are organizing their collections for the fashion show.  
The stylist confirmed that the designers have organized their collections for the fashion show.

The family noticed that the gardeners are planting the trees in the backyard.  
The family noticed that the gardeners have planted the trees in the backyard.

The teacher thinks that the tutors are helping the students with their homework.  
The teacher thinks that the tutors have helped the students with their homework.

The teacher supposes that the students are checking their email in the library.  
The teacher supposes that the students have checked their email in the library.

The nun thinks that the tourists are dirtying the cathedral during their visit.  
The nun thinks that the tourists have dirtied the cathedral during their visit.

The bride believes that the coordinators are arranging the seating for the wedding reception.  
The bride believes that the coordinators have arranged the seating for the wedding reception.

The organizer believes that the volunteers are finishing the preparations for the festival.  
The organizer believes that the volunteers have finished the preparations for the festival.

The media confirmed that the actors are donating the money to charity.  
The media confirmed that the actors have donated the money to charity.

The professor announced that the editors are approving her article for the journal.  
The professor announced that the editors have approved her article for the journal.

The radio announced that the guitarists are organizing a concert in the convention center.  
The radio announced that the guitarists have organized a concert in the convention center.

The reporter affirmed that the scientists are testing the vaccine on rats.  
The reporter affirmed that the scientists have tested the vaccine on rats.
The magazine indicated that the actors are answering the letters from their fans. The magazine indicated that the actors have answered the letters from their fans.

The secretary said that the assistants are accusing their boss of fraud. The secretary said that the assistants have accused their boss of fraud.

The secretary noticed that the assistants are removing the trash from the office. The secretary noticed that the assistants have removed the trash from the office.

The secretary thinks that the assistants are dividing the files into three piles. The secretary thinks that the assistants have divided the files into three piles.

The superintendent guarantees that the instructors are testing the students appropriately. The superintendent guarantees that the instructors have tested the students appropriately.

The designers noticed that the models are practicing their walk on the runway. The designers noticed that the models have practiced their walk on the runway.

The students noticed that the professors are placing their quizzes on the bookshelf. The students noticed that the professors have placed their quizzes on the bookshelf.

The tenants noticed that the electricians are fixing the powerlines in the building. The tenants noticed that the electricians have fixed the powerlines in the building.

The children think that the magicians are turning the rabbits into doves. The children think that the magicians have turned the rabbits into doves.

My neighbor noticed that the students are watching the game on TV. My neighbor noticed that the students have watched the game on TV.
VITA

Rosa E. Guzzardo Tamargo

Education
Ph.D. in Spanish (Hispanic Linguistics) 2012
Dual title in Language Science
The Pennsylvania State University, University Park, PA

M.A. in Spanish 2008
The Pennsylvania State University, University Park, PA

B.A. in Hispanic Studies 2006
Double-major in Modern Languages
Universidad de Puerto Rico, Río Piedras, PR

Publications


Presentations (selection)


Teaching experience
SPAN 497 – Puerto Rican Culture
SPAN 410 – Advanced Oral Expression
SPAN 215 – Introduction to Hispanic Linguistics
SPAN 110 – Intermediate Conversation
SPAN 003 – Intermediate Spanish

Other professional experience
Mentor for PSU undergraduate research assistants 2009-2012
Organizing Committee Member for PSU Center for Language Science 2010-2011
Graduate Student Young Scholar Speaker Series
Assistant Director of PSU Summer Education Abroad Program in Puerto Rico 2008-2011