Grammatical Optionality and Variability in Bilingualism: How Spanish-English bilinguals limit clitic-climbing

by

Danielle L. Thomopoulos Thomas

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Spanish and Portuguese
University of Toronto

© Copyright by Danielle L. Thomopoulos Thomas 2012
Grammatical Optionality and Variability in Bilingualism: How Spanish-English bilinguals limit clitic-climbing

Danielle L. Thomopoulos Thomas
Doctor of Philosophy
Department of Spanish and Portuguese
University of Toronto
2012

Abstract

This thesis considers how different groups of Spanish speakers (monolinguals, early bilinguals and late bilinguals) organize and limit grammatical optionality related to the placement of Spanish pronominal clitics with many complex infinitival constructions (Spanish clitic-climbing). In examining empirical work on the process and outcome of early and late dual language exposure and how early and late bilinguals acquire and limit grammatical optionality, this study will contribute to our understanding of 1) the nature of language-related cognition at different ages; 2) the systematic nature of bilingual language behaviour in child and adults (transfer, cross-language influence, etc.); 3) the cognitive and contextual factors associated with age of exposure to bilingualism to explain bilingual language behaviour; and 4) the importance of incorporating a clear model of language variation (language-internally and cross-linguistically) into a formal model of (bilingual) language.

The empirical study conducted here tested how highly proficient heritage speakers (HS) of Spanish (native speakers of Spanish and Spanish-English bilinguals) deal with the optionality of clitic-climbing structures compared to monolingual speakers (native speakers) and highly proficient adult L2 speakers of Spanish (Spanish-English bilinguals). Forty participants completed a picture elicitation task testing a lexical limitation of the optionality, and an
acceptability-preference task testing the speakers’ judgments on structural, semantic and lexical limitations of the optionality. Results show that all groups of speakers exhibited knowledge of syntactic constraints associated with pronominal placement in Spanish (optional clitic-climbing) infinitival sentences. All groups also performed similarly in exhibiting sensitivity to non-categorical factors that have been shown to guide the preferences of monolingual Spanish speakers. However, in the production task, the heritage speakers significantly outperformed the monolingual and non-native speakers of Spanish in their use of the Spanish-specific variant (proclisis). I explain these results through both cognitive and contextual factors related to age of exposure to bilingualism, and I discuss how the production results may underestimate a monolingual-bilingual difference for this optional domain.
“Si elle avait pu demeurer éternellement à la frontière des choses, elle aurait été parfaitement heureuse. La plupart des gens lui semblaient assoiffés par les résultats, tandis que elle se contentait de prospectives...” —Agnes Desarthe

This is for my parents, Diana and Danny, the long-awaited result of their love, extraordinary patience and back-breaking hard work...

...and that of Baba Daphne, Baba Elsie, Dedo Dinko and others who I don’t even know about, but who had to work a lot harder than me to get me here.
Acknowledgments

This dissertation is the result of the sum of the influences I have had through my education, work, and being in certain places at certain times. Some didn't know they were having as big an impact as they did, while others should not be surprised to find themselves here. This list of acknowledgments is long, but the path has been long as well and each person here deserves to be mentioned.

First, I thank those who had the earliest, but most long-lasting influence: Tanya (who taught me presentational skills—how to talk to people!), Mr. Procter (who taught me the value of writing well), Megan Buckley (who I learned about great leadership through watching), Professor Vercollier (who taught me professional elegance and personal responsibility in the classroom), Professor Bartlett (who taught me about the fun of learning and professional initiative), Melanie (who taught me what real compassion is), Christiane (who taught me how to set personal boundaries), Rosalin (who taught me that there are no boundaries) and Clelia (who taught me to be brave, even if it is unpopular).

Second, I am grateful to those friends who come and go from the radar with the comings and goings of life, but who I love and appreciate for all their support: Tanya, Alex Cragg, Christiane, Chris, Rosalin, Shanna, Adam, Isabelle B., Sharon, Lanette…and all others.

Third, for the great groups of people who have taught me, supported me, encouraged me and been a great influence on my personal and professional growth:

The Champion group: Reza Ghasry, Dr. David and Jeanette Lane and family, Dr. Joe and his family, Lanette Ward, Donna, Mark and Elliot Morrison-Reed, Catherine and Paul Marsiglio and family, Dr. Freda Miller, Jennifer Robertson-Decker, Dave Menikefs and especially Sharon Mah and all her family. I am so lucky to have known such an extraordinary group of human beings. I thank them for their kindness, generosity and being a great example that the terms —high-achieving‖ and —good people‖ are compatible!

For those at Brock, UTM and York who gave me incredible opportunities in supportive environments to work and become a better teacher: Irene Blayer, Cristina Santos, Cora Solway, Emmanuel Nikiema, Michael Lettieri, Rosa Ciantar, Belinda Grayburn, Gabriela Alboiu, Peter Avery, Josie Sansonetti, Kasia Wolf, Marina Smeriglio and Silvana DeBona.

To the Spanish and Linguistics Departments at U of T, many thanks for the work opportunities, and to the phenomenal teachers I’ve had and people who have helped me along the way: Stephen Rupp, Josiah Blackmore, Bob Davidson, Victor Rivas, Mike Barrie, Alanna Johns, Blanca Talesnik, Rosinda Raposo, and of course Ana, Laura and Cristina, but more about them later.

To the Linguistics group in the Spanish department, thanks for their advice, information, complaint forum and potlucks: Irina, Anna L, Natalia M, Yasaman, Yadira, Tanya, Joanne and Olivia.
For those in the object omission group: Yves Roberge, Ana Pérez-Leroux, Mihaela Pirvulescu, Isabelle Belzil, and all those who have come and gone from the group—it is here that I learned how to be a great academic, both professionally and personally. The value of this experience—of just watching greatness in action—is immeasurable and I am forever grateful to Yves, Ana and Mihaela for including me unconditionally all these years.

To the Thursday night crowd: Clelia, Isabel, Natalia—us the original four and all those others who have come to join us through the years—Arturo, Agnes, Ramiro and Raúl, Nelson, Martha, Sylvia, Anna C., Diana, etc.

And finally, my Salvadorean family: Clelia, Rodolfo, Carolina, Coni, Carlos (and all the extended Salvadorean group)—I can not ever express how grateful I am for having them include me at a time when one family wasn‘t enough. Along with my own family, they helped me stay afloat, fed me, listened to me, and rescued me over and over again.

And to those whose influence resulted in the most direct impact on my career path to this point: For Professor Portebois: Who taught me how to think, be enthusiastic and kind in the classroom, tackle a big project and not get overwhelmed, and how to be confident about my abilities—I only continued on to graduate school because she told me I could—thank you!

For Diana: I‘ve said it before, but there is no one better than I could think of to have been my partner year by year in this. I thank her for her kindness, availability whenever I needed her voice in the recording studio (for tesina and for thesis) and helping me study for comps when I had no idea what I was talking about!

For Alejandro: Who paved the way for me like a brother, and included me in his family like a sister. I‘m always grateful for his help, encouragement, trust and love. I look forward to a great career seeing if I can keep up with him!

For Melanie: The person with the most—mosts of anyone I know—the person with the most compassion, the most sensitivity, the most loyalty, the most beauty and who is the (most) best friend to anyone who is lucky enough to call her a friend. I thank her for understanding me and for being there when I have needed her!

For Clelia: From bloqueo! to Nnanga!, I am a lucky person to have gone to Cuba and then surrendered to the ride. Here, she had the painstaking task of editing all the Spanish work and providing countless hours of encouragement to just get it done! She has pushed me in both my professional and personal lives, and I am always grateful for her trust in allowing me to know her past, her daughter, her family, her work and her plans for the future.

And to my committee, who is more than just a committee:

For Cristina: Who is the angel on my shoulder whispering how to be a great linguist and even better teacher. One meeting always changes my life in some way—thank you!
For Laura: I really don’t think I could ever express what she has meant for the path of my life—she changed the course of my life with only one course! Her calmness and kindness are unmatched, and I am forever indebted for what she (quietly) did and continues to do for me. Thank you!

For Ana: I’m so glad I got it and decided to trust her when she told me to—I am a better person for it both professionally and personally. She has been the greatest teacher and mentor—unconditional, knowledgeable, encouraging, forthcoming and patient. I will never forget this and will pass it on in my own career. Thank you! Many thanks as well to Bill, Michel and Paul for welcoming me at home so often!

AND finally to my family:

For aunts, uncles, cousins, everyone: thanks for their encouragement, support, love.

For Daphne and Alfie: I will always proudly wear my ring and call myself the ―other wife‖. I am so grateful for their including me, letting me be DD and taking care of me when I’ve always needed it. And to Alfie, for all the technical expertise and sound work he did for nothing—thank you, thank you!

For Jack and Roxanne: Who have provided me with more moments of joy than they will ever know about, and have been the one thing that I could count on to make me happy and make me forget about the work. I love you!

For my parents, Diana and Danny: who have provided me with all that I’ve ever needed, gave me the most incredible childhood, supported me through a difficult adulthood and saved my life when it needed saving. The best traits of each of them are lofty heights to aspire to, but this is what I hope to be able to do in my career and in my life in general: be as compassionate, giving and tireless as my mother and as generous, sensitive and hard-working as my father. There is not one aspect of their parenting I would change—they are the model for me of how parents should be and it will be hard for me to find a life of my own as great as the one they have given! Despite my training as a linguist, I have no words in any language to express the degree of my gratitude, awe and love for them as parents and human beings. Put simply, thank you.
# Table of Contents

ABSTRACT ................................................................. ii

ACKNOWLEDGEMENTS .................................................. v

TABLE OF CONTENTS ................................................... viii

LIST OF TABLES ........................................................ xii

LIST OF FIGURES ....................................................... xiii

LIST OF APPENDICES ................................................... xiv

**Chapter 1. Introduction** ........................................... 1

1.1 Language basics .................................................. 1

1.2 The current study ............................................... 4

1.3 Empirical scope and organization of thesis .................... 11

**Chapter 2. Language Theory** ................................... 15

2.1 Introduction ...................................................... 15

2.2 Language and cognition ......................................... 15

2.2.1 Language and Minimalism .................................. 15

2.2.2 Language performance ....................................... 20

2.2.3 Theories of language variation ............................. 26

2.2.4 A model of language representation and use ............ 30

2.3 Pronouns and infinitival predicates in Spanish and English .... 32

2.3.1 Variation and optionality in the Spanish pronominal system ... 32

2.3.2 Theories of pronouns and Romance clitics ................ 38

2.3.2.1 Movement approaches .................................. 41

2.3.2.2 Base-generation approaches ........................... 43

2.3.3 Theories of complex predicates & clitic-climbing .......... 45

2.3.4 The learning task for Spanish-English bilinguals .......... 54

2.4 Chapter Summary ................................................ 56
Chapter 3. The Language of Bilingualism ................................................................. 57
  3.1 Introduction ........................................................................................................ 57
  3.2 Language and bilingualism ................................................................................. 57
    3.2.1 Language autonomy and interdependence ................................................. 57
    3.2.2 Functional and contextual bilingual effects .............................................. 59
    3.2.3 Language-internal factors of bilingual language behaviour .................... 62
    3.2.4 Language-external factors of bilingual language behaviour .................... 70
    3.2.5 Variability and interpretability in bilingualism .......................................... 77
  3.3 A systematic approach to bilingual language behaviour ................................. 82
    3.3.1 A bilingual language system ..................................................................... 83
      3.3.1.1 The bilingual lexicon ........................................................................ 83
      3.3.1.2 Language selectivity in bilingualism ............................................... 87
      3.3.1.3 A language model for bilingual representation and use ................. 90
    3.3.2 A typology of bilingual language behaviour (single-language) ............ 96
      3.3.2.1 A systematic view of quantitative and qualitative effects ................. 97
      3.3.2.2 General bilingual effect and cross-language influence .................... 99
      3.3.2.3 Transfer ........................................................................................... 106
  3.4 Chapter Summary .............................................................................................. 109

Chapter 4. Early and late bilingual development and attainment ......................... 111
  4.1 Introduction ....................................................................................................... 111
  4.2 Age of exposure in early and late bilingualism .............................................. 112
    4.2.1 Age and a critical period for bilingual acquisition .............................. 112
    4.2.2 Age, exposure and the context of learning ............................................ 116
      4.2.2.1 Context of exposure in early and late bilingualism ....................... 117
      4.2.2.2 Exposure in heritage bilingualism ............................................... 118
      4.2.2.3 Exposure in adult L2 acquisition ............................................... 120
      4.2.2.4 Naturalistic vs. classroom learning ............................................ 123
  4.3 Development and attainment in early and late bilingualism ....................... 126
    4.3.1 Simultaneous and heritage bilingualism ............................................... 127
      4.3.1.1 Early childhood ......................................................................... 127
      4.3.1.2 Mid-late childhood & adolescence ............................................. 131
4.3.1.3 Linguistic outcomes of heritage bilingualism .................133
4.3.2 Adult L2 acquisition ..............................................................136
  4.3.2.1 Initial state and process of adult L2 development ............136
  4.3.2.2 Linguistic outcomes of adult L2 and for adult L2 Spanish
            pronominal system ......................................................140
4.4 Chapter Summary ...............................................................142

Chapter 5. The Study: research questions, methods & hypotheses ..............143
5.1 Introduction .............................................................................143
5.2 Research questions .................................................................143
5.3 Methodology ............................................................................146
  5.3.1 Participants ........................................................................146
  5.3.2 Structures under analysis .....................................................150
  5.3.3 Tasks and testing session .....................................................153
    5.3.3.1 Picture elicitation task ..................................................153
    5.3.3.2 Acceptability-Preference task ........................................156
5.4 Hypotheses of study ...............................................................159

Chapter 6. Results ...........................................................................164
6.1 Syntactic knowledge of verb incorporation ..................................164
6.2 Lexical limitations on proclisis ...................................................169
  6.2.1 Picture elicitation task .........................................................169
  6.2.2 Acceptability-Preference task .............................................178
  6.2.3 Lexical constraints and task-type ..........................................182
6.3 Semantic limitations on proclisis ................................................187
6.4 Summary of results .................................................................193

Chapter 7. General Discussion & Conclusions ....................................194
7.1 Introduction ..............................................................................194
7.2 Knowledge & use of clitic-climbing by Spanish speakers ..............194
7.3 Age of exposure: language & cognition ......................................201
7.4 Age of exposure: context of exposure .......................................206
7.5 Optionality and variability in bilingualism: concluding remarks........208
7.6 Directions for future research .................................................................210

References .......................................................................................................212
List of Tables

Table 2.1: Frequency counts and proclisis by verb type in Spanish speech corpus, Davies (1995) .................................................................36

Table 5.1: Criteria for selection of participants in current study ........................................149

Table 5.2: Conditions used to test Verb Incorporation in Spanish ..................................151

Table 5.3: Hypotheses for structural, lexical and semantic variables for the current study ......163

Table 6.1: Acceptability of proclisis by group and verb type ........................................165

Table 6.2: Acceptability of proclisis by group and embedded negation .........................166

Table 6.3: Percent of direct object responses by group and verb type (production task) ......169

Table 6.4: Percent use of proclisis by verb type and group for pronominal responses ........171

Table 6.5: Average rate of acceptability of proclisis by verb type and group ..................178

Table 6.6: Preference of pronominal word order options by verb type and group ..........180

Table 6.7: Average rate of acceptability/preference and production of proclisis by verb type and group (task-type comparison) ........................................183

Table 6.8: Statistical Analyses of limitations of proclisis by task type, verb type and group ...183

Table 6.9: Average rate acceptability of proclisis by tense of ir a/interpretation of predicate and group ..........................................................188

Table 6.10: Preference of pronominal word order by tense of ir a/interpretation of predicate and group ..........................................................190

Table 7.1: Summary of results testing syntactic knowledge of verb incorporation ...............195

Table 7.2: Summary of results for use of lexical constraints on limiting optionality in production and receptive tasks ..........................................................197

Table 7.3: Summary of results for use of tense of ir a/interpretational constraints on limiting optionality in a receptive task ..........................................................199
List of Figures

Figure 2.1: The Language Component of a Simultaneous Bilingual .................................18
Figure 2.2: Language Model for Production ........................................................................32
Figure 3.1: A Language Model for Bilingual Speakers .........................................................95
Figure 6.1: Acceptability of proclisis by group and embedded negation ......................164
Figure 6.2: Monolingual responses (individual results embedded negation) .............167
Figure 6.3: Heritage Speaker (HS) responses (individual results embedded negation) ...167
Figure 6.4: L2 Responses (individual results embedded negation) .................................168
Figure 6.5: Monolinguals ir a (individual results for use of proclisis) ..............................172
Figure 6.6: L2 speakers ir a (individual results for use of proclisis) .................................172
Figure 6.7: Monolinguals querer (individual results for use of proclisis) .....................173
Figure 6.8: L2 speakers querer (individual results for use of proclisis) .........................173
Figure 6.9: Average rate of proclisis in production by verb type and group ..................176
Figure 6.10: Acceptability/Preference of proclisis across verb types and group ..............181
Figure 6.11: Proclisis by task type (and group)—IR A ......................................................184
Figure 6.12: Proclisis by task type (and group)—QUERER ............................................184
Figure 6.13: Proclisis by task type (and group)—PREFERIR .......................................184
Figure 6.14: Average rate of proclisis by task type for group and verb type ...............187
List of Appendices

Appendix A: Davies’s (1995) results for proclisis by verb type and mode of production
Appendix B: Language History Questionnaire/Interview (Spanish & English)
Appendix C: Summary of Participants’ Profile
Appendix D: Rater Assessment for Narrative Data (Spanish and English)
Appendix E: Instructions for Acceptability/Preference Task and Picture Elicitation Task
Appendix F: Acceptability/Preference Task (APT): Version A
Appendix G: Sample Test Sheet for APT: Version A
Appendix H: Picture Elicitation Task—Contexts and Test Questions
1 Introduction

1.1 Language basics

Being able to communicate linguistically involves the integration of a complex set of procedures related to two fundamental tasks: being able to analyze and interpret what others say and being able to formulate speech that is interpretable. In the first place, one must develop knowledge of the sounds, forms and structures of the language that is being used in order to assign an interpretation to the utterance for comprehension or production. In the second place, one must have a process by which one takes an intended interpretation and transforms it into interpretable speech (and vice-versa). In addition, one must develop knowledge of the different ways that a particular interpretation can be uttered, depending on the context in which, and with whom, it is being used. These three components represent the linguistic, psycholinguistic and sociolinguistic knowledge that anyone engaged in normal patterns of linguistic communication is responsible for.

To complicate the matter, this system must also be able to accommodate the developmental and synchronic aspects of language contact situations. For early bilinguals,\(^1\) the system must be able to account for evidence of both early grammatical autonomy as well as bilingual interaction effects in development and in mature grammars (code-switching, transfer, cross-language influence, etc). For late bilinguals,\(^2\) the system must be able to account for the increased or decreased variability—where variability is defined as developmental or mature linguistic patterns that have not been attested in monolinguals—that often characterizes second language (L2) knowledge and use. In modelling a language system, de Bot (1992) highlighted the importance of accommodating contexts of bilingualism and the three components presented above—linguistic, psycholinguistic and sociolinguistic knowledge (italics are mine):

\[^1\text{Here, the reference to ‘early bilinguals’ will include speakers who had very early exposure to two languages in childhood (<4;0): either simultaneous (2L1) or early child sequential (L2) bilingualism.}\]

\[^2\text{Here, the reference of ‘late bilinguals’ will include speakers who had their first exposure to a second-language (L2) as teenagers (>15 years old) who would have already arrived at the onset of puberty.}\]
“In language behaviour research there have traditionally been reasonably sharp dividing lines between linguistic, psycholinguistic and sociolinguistic research. In a good production model [language model] these dividing lines fade; the model should be able to cope with universal characteristics of language as well as cognitive processes and situational factors in interaction and their consequences for language use.” (p. 1-2)

This quote above alludes to there being a number of fields of study that form the different components of a large linguistic spectrum. The division of this spectrum into its modern component parts began fundamentally with the Chomskyian era in the 1950s. In describing the fundamental difference between competence and performance, Chomsky (1965) highlighted the need to go beyond the descriptive aims of the previous behaviourist era in order to explain for the complexity and regularity of the system regardless of any one speaker’s experience with their language environment. This system appeared so extraordinary that many researchers of the modern linguistic era proposed that, like other biologically-regulated and specialized cognitive systems, the linguistic system must be as specialized in a way so as to mediate the representational and developmental demands that the generalized cognitive system couldn’t possibly support (Lenneberg 1967; Pinker 1984; Chomsky 1993). Within the scope of the generative framework for linguistic studies, two basic questions were identified as the logical and developmental problems of language: 1) How does a speaker know what they know about language with such impoverished input?; and 2) How do speakers of the same language get from having no knowledge of that language to having complete knowledge of the language in such a uniform way and with relative ease?

Where specific linguistic knowledge is concerned (constraints on sounds, forms and structure), this study will adopt a generative framework and assume that there is a language-specific cognitive component that provides an inventory of universal linguistic features (i.e. Universal Grammar (UG)) and two mechanisms to regulate the constraints specific to any one language: 1) a lexicon where a language-specific inventory of lexical and functional items are established through exposure to input and stored for use in comprehension and production; and

---

3 Saussure (1916) proposed the fundamental distinction between language knowledge and use as *langue* and *parole*. This distinction was captured in the early generative proposals of Chomsky as competence and performance.
2) a computational system that associates these lexical and functional items according to language-specific lexical specifications and universal constraints on structure-building, and in turn, which allows a speaker to provide an interpretation (in comprehension and production) for an infinite set of utterances from their experience with a finite set of utterances in the input. In terms of development, the current study will not examine or compare speakers at different moments of the acquisition process, and as such, no particular developmental questions are addressed here. This study focuses instead on the notion of complete knowledge and the role of input in attainment: What do we mean by complete knowledge in generative studies and can it accommodate both single- and dual-language contexts? And, to what degree can the quantity and quality of input explain the unique linguistic behaviours of speakers who have been exposed to bilingualism at different ages?

Many researchers have been occupied over the past number of decades in answering questions related to the processing of language in comprehension and production, and how speakers acquire and utilize information related to the linguistic and situational contexts of a speech act in order to limit the variety of linguistic options available to them. In both the fields of psycho- and sociolinguistics, a variety of proposals have been put forth to account for performance-related aspects of the choice (i.e. selection) involved in speech production. In the current study, the model of speech production proposed by Levelt (1989) will be used as the basis for the performance-related aspects of language knowledge. This system is preferred over others because it is compatible with the lexicalist nature of a Minimalist grammar, and it provides a systematic framework that can help explain many of the bilingual language behaviours exhibited by different groups of bilingual speakers. Chapter 2 describes the details of this system and the various proposals related to the way that grammatical variation (different registers, etc.) can be accommodated in a language system.

1.2 The Current Study

Empirical evidence (of language knowledge and production) must be integrated into a simple and efficient language model that can accommodate both the variety of lexical, phonological and morphosyntactic forms that exist within and across languages, and a number of constraining variables. The goal of the current study, therefore, is to not only examine a detailed area of grammar in order to better understand what aspects of the linguistic component of the overall
language system are available and/or utilized by different types of speakers (i.e. monolinguals and bilinguals), but also to consider how the linguistic, psycholinguistic and sociolinguistic tendencies of both monolingual and bilingual speakers can be accommodated in one simple and efficient language system.

In general, one of the most challenging aspects to account for in a simple and efficient language model is grammatical optionality because it implies acquiring both linguistic knowledge (what the underlying grammar of the options looks like) and performance-related knowledge (how those options are chosen in everyday communication). Here, a distinction is made between optionality and variability: i) optionality refers to a speaker’s *competence*—that is to say, the knowledge a speaker has of the morphosyntax that underlies the structural options available for the same utterance; and ii) variability refers to the *performance* of any particular speaker/group of speakers regarding the choice of those structural options. In examining variability, a variety of language-internal and language-external factors may affect the way that any one speaker/group of speakers chooses one structural variant over the other (in production or in relation to judgments of preference). The aim here is to examine the nature of optionality (*linguistic competence*) in the grammars of bilingual speakers and to begin to determine which factors (language-internal and/or external) bilingual speakers use to make language choices (*language performance*), and as such, to limit random optionality in their grammar.

The notion of *variability* described above is slightly different from the way it is often used in the study of second language. In the field of L2 acquisition, *variability* is often meant to refer to the non-target-like ways that late learners use certain grammatical domains during development and/or at the most advanced stages of proficiency (Sorace, 1993, 2005; White, 2003). For example, advanced L2 speakers of English often exhibit variable patterns of use for the third person present tense marking (-s), as in (1), compared to English monolinguals, who do not exhibit such variability, as in (2) (e.g. Lardiere, 2000). Native English speakers do not exhibit variability in production for this domain; the use of the present tense morpheme is obligatory with third person singular subjects. Adult bilinguals who variably omit the third person morpheme (individual speakers or within the group) may do so for different reasons related to language-related cognitive functioning (representational or processing issues) and/or language experience (quality of exposure, etc.).
The difference between the type of variability exemplified above for present-tense marking and the type of variability in question in the current study is that, unlike in the case of English present-tense marking, for Spanish clitic-climbing both forms are target-like. As such, the aim of the current study is to not only examine the linguistic knowledge that Spanish speakers (monolinguals and bilinguals) have of the optionality in question, but also to consider the ways in which different groups of speakers (monolinguals/bilinguals; children/adults) organize and limit a random grammar depending on their experience with the language and/or cognitive abilities. Even though all speakers are expected to exhibit some variability for such a grammatical domain, the question for monolinguals and bilinguals is whether the patterns of variability are the same, and if not, how their age or context of exposure to the language can explain the different ways that these speakers exhibit knowledge and/or make choices in production for this domain.

Given the notions of optionality and variability outlined above, the aim here will be to examine the degree to which monolingual and bilingual speakers have knowledge of grammatical optionality (and its linguistic constraints) and limit their preferences for the competing structural options according to language-internal factors. More specifically, the aim will be to determine if Spanish monolinguals and Spanish-English bilinguals i) have knowledge of the grammatical optionality (and constraints) that characterizes the positioning of pronominal objects with complex infinitival predicates, illustrated in (5); and ii) to determine the extent to which these speakers limit a random grammar by attesting to the lexical type of matrix verb and/or the semantics of the matrix verb.

(5) a. Juan quiere verla. / Juan puede verla.  
    John wants to see it. / John can see it.

b. Juan la quiere ver. / Juan la puede ver.  
    *John it wants to see. / *John it can see.

Compared to English, the Spanish sentences in (5) have been shown to differ in two linguistic ways: 1) Spanish object pronouns are clitics (Xo) while English object pronouns have the same grammatical status as lexical objects (XP) (Kayne, 1975; Strozer, 1976); and 2) with a variety of subject control and subject-raising predicates, Spanish pronominal clitics can surface in
the unmarked position (like in English), right-adjacent to the infinitive that has selected the argument as in (5a), or in the language-specific position, left-adjacent to the matrix verb of the complex predicate as in (5b). Due to the apparent way that the clitic has “climbed” up to the higher clause, the language-specific variant is often referred to as “clitic-climbing” (Kayne 1989). In the present study the term ‘clitic-climbing’ will refer to these particular constructions that instantiate both unmarked post-verbal (enclisis) and language-specific pre-verbal (proclisis) structural variants.

While the optionality in (5) appears i) to be a word order issue, and ii) to be a case of random variation, a considerable amount of theoretical and sociolinguistic work has been done to determine that:

i) the word order difference in (5) is due to an underlying structural difference related to the complex predicate, and

ii) the optionality in (5) is limited by a variety of language-internal factors.

In the first case, evidence of categorical constraints operating to block proclisis, as in (6a) and (6b), led to a variety of proposals suggesting that in the case of proclisis (repeated in (6c)), the two verbs of the complex predicate were reduced (or restructured) to form a mono-clausal configuration, while in the case of enclisis, the two verbs of the complex predicate formed a bi-clausal configuration (Aissen & Perlmutter, 1976; Rizzi, 1978; Moore, 1996; Masullo, 2004). The surfacing of the clitic in both cases follows language-specific rules on pronominal word order with tensed and untensed predicates: in the case of the reduced clause (see (5b)), the clitic identifies the overall predicate as tensed (pron-(V-V)), and in the case of the non-reduced clause (see (5a)), the clitic identifies the lower predicate as untensed (infinitive-pron).

(6)  

a. *Juan [la [quiere que Ana vea]]. / Juan [quiere [que Ana la vea].
   John it wants that Ana sees / John wants that Ana it sees
   “John wants Ana to see it.”

b. *Juan [la [quisiera no saber]]. / Juan [quisiera [no saberla].
   John it would like not to know / John would like not to know it
   “John would like to not know it.”

c. Juan [la [quiere ver]]. / Juan [quiere [verla]].
   John it wants to see / John wants to see it
   “John wants to see it”

Masullo (2004) proposed a Minimalist account of the two structural options in (6c) that places the optionality as an order of merge operations: merge of the infinitive and the pronominal clitic
first (unmarked option, enclisis), or merge of the matrix verb and the infinitive first (languagespecific option: verb incorporation, proclisis). This proposal will be used as the basis for the syntactic approach to clitic-climbing in the current study and described further in Chapter 2.

This optionality has also been examined from a sociolinguistic perspective. In general, two language-internal factors related to the complex predicate have been proposed to act as limiting factors of the optionality in (5): the type of matrix verb (lexical variable) and the relative interpretation of the matrix verb (semantic variable). Following research previously conducted on written production in Spaulding (1927, cf. Davies, 1995), Keniston (1937, cf. Davies, 1995), Myhill (1988) and Rosen (1989, cf. Davies 1995) among others, Davies (1995) conducted a large-scale corpus study of monolingual production (speech and writing) to determine which factors monolinguals used to limit the word order optionality of clitic-climbing constructions. While a variety of factors were found to limit the optionality (multiple clitics, reflexivity of clitic, animacy of clitic, etc.), the type of matrix verb was found to be one of the most prominent determinants. While described as a lexical variable in the current study, Davies noted that differential rates of proclisis appear to occur across a continuum of verb types related to the relative semantic weight of the matrix verb. He also noted three previous characterizations of this continuum: how auxiliary-like the matrix verb is (Napoli, 1982), how grammaticalized the matrix verb is (Myhill, 1988), and whether the matrix verb is light or heavy (Rosen, 1989, cf. Davies, 1995). Regardless of the characterization, Davies basically confirms this semantic-based analysis: of all the verbs Davies analyzes here (32 in total), the ones with the highest rates of proclisis “nearly all[…]fall into the auxiliary/aspectual/modal set of verbs” (p. 373). Nevertheless, there is no analysis of the varying light and heavy semantic interpretations that a verb like ir a, ‘to be going to’, can have (future or motion interpretation), nor is there an account of why verbs that don’t fall into the auxiliary/aspectual/modal set of verbs also exhibit relatively high rates of proclisis (i.e. saber, ‘to know’); consequently, in the current study, the type of matrix verb, as a

---

4 In Davies’s study, monolinguals use proclisis with ir a in speech production 86% of the time and in writing 66% of the time. Davies includes ir a in the auxiliary/aspectual/modal class of verbs, but doesn’t make any note that he restricted his analysis to the cases when ir a was used with a “light” interpretation. There appears to be an implicit assumption here, as in other studies, that ir a + infinitive is used to express the Spanish future tense.
linguistic variable, is identified as a lexical factor that limits the grammatical optionality characteristic of this domain.

To address the question of the semantic variable—whether the relative semantic weight of the matrix verb limits the rate of proclisis—an analysis is, thus, needed where the same matrix verb differs in interpretation. While there are many examples of this type of interpretational duality in Spanish (future vs. motion *ir a* ‘to be going to’, deontic vs. epistemic *poder* ‘can/to be able to’, wish vs. desire *querer* ‘to want to’), no study has examined this question empirically for any group of Spanish speakers. Napoli (1982) offered a descriptive account of how proclisis is limited in Italian (another clitic-climbing language) by using a range of desiderative verbs, (*volere* ‘to want/wish for’ > *desiderare* ‘to desire’ > *anelare* ‘to yearn’), but again no empirical analysis was provided to determine if the optionality is limited when the same matrix verb varies in its interpretation across the light-heavy continuum. The semantic question, then, remains open: If proclisis is limited when the type of matrix verb used has a more strong lexical interpretation, then shouldn’t proclisis be more limited when the same matrix verb is used with a heavy lexical interpretation, as compared to a light auxiliary-like interpretation?

In considering the ways in which clitic-climbing constructions have been previously analyzed from theoretical and variationist perspectives, a number of questions arise for the current study that fall under both generative and more performance-related frameworks. In the first place, acquiring the optionality above implies dealing with a challenging learnability problem: How does any one speaker come to know the underlying representation of this domain, and the categorical constraints of the optionality by virtue of input alone? And more specifically, can non-native and/or bilingual speakers acquire the syntactic constraints of this grammatical domain in a way similar to monolingual speakers? Secondly, dealing with optionality implies making a choice, and as such, both psycholinguistic and sociolinguistic factors will affect any one speaker’s word order choice during production. I pursue the following questions here:

1) How does a bilingual speaker’s experience with an unmarked grammatical representation affect their activation and selection of competing structural options?

2) Are bilingual speakers sensitive to the non-categorical linguistic constraints that have been shown to limit the optionality in monolinguals?
3) What cognitive and contextual factors related to age of exposure to Spanish may account for the different ways in which fluent bilinguals and monolinguals analyze the optionality and use the competing word orders? Here, by examining the core linguistic knowledge that bilingual speakers have (or don’t have) of the syntactic constraints on competing syntactic variants for the same grammatical domain, we may be able to offer a better systematic account of how competing structural variants are represented in a formal language model.

To answer these questions, I examine three groups of Spanish speakers in this study: monolinguals, heritage speakers of advanced Spanish proficiency and adult L2 speakers of advanced Spanish proficiency. Here, I use the term ‘heritage speaker’ in a linguistic sense (see Valdés, 2000, Carreira, 2004); these speakers were born and educated in an English-language community, but spoke Spanish consistently throughout childhood and adolescence as a home language with native Spanish-speaking parents (adult immigrants). Further, while there is some debate over whether heritage speakers are simultaneous (2L1) or early sequential (child L2) bilinguals (Spanish L1-English L2) (see Montrul, 2008), I assume that the heritage speakers examined in the current study had a language experience that is most closely associated to simultaneous bilingualism. This assumption is made on the basis of two factors: 1) that all heritage speaker participants identified having been exposed to English on a consistent basis before 2;0 (daycare, community contact, neighbours and peer groups); and 2) that Toronto, where all of the heritage speakers were born and raised, has a relatively small population of Spanish speakers (census 2006: 2%) and relatively disperse Hispanic communities (as compared to more dense communities in the United States). Further, like other characterizations of heritage speakers (see review in Montrul, 2008), I will also assume that these heritage speakers experienced a shift in dominant patterns of exposure around the age of 4;0, from Spanish to English; consequently, these speakers were mostly Spanish-dominant (vocabulary size, overall fluency) before starting kindergarten and then became more English-dominant (vocabulary/morphosyntactic measures, overall fluency and literacy, etc.) after beginning full-time school. The effects that these patterns of exposure can have on the development and outcome of this type of bilingualism are discussed in more detail in Chapter 4.
Finally, the reasons for including both heritage speakers and L2 speakers in the same study are twofold. First, the genesis of this project occurred through considering an oft-repeated notion in research on adult L2 ultimate attainment. In general, L2 speakers typically “fail to acquire the target language grammar” (Birdsong, 1992, p. 706), where the target language grammar is understood to be that represented by monolingual (native) speakers. Although researchers engaged in the study of acquisition understand that a variety of cognitive, contextual and socio-affective factors may be the cause of any one speaker’s inability to reach so-called complete outcomes, the way that this statement reads to most others (teachers, language learners, other scholars, etc.) is that adults are somehow deficient when it comes to language learning. While it may be true that late adult learners don’t typically achieve the same grammatical outcomes of those who have acquired a language from birth as a monolingual speaker (so-called complete outcomes), this should not be surprising given that the overall objective of L2 acquisition is not to become a monolingual speaker, but rather, a bilingual one. Consequently, the basic question that we must address in the L2 field before making any claims about an individual’s ability or lack of ability to acquire a language as an adult is the following: What is a complete outcome for a bilingual speaker? As such, examining simultaneous bilingual speakers can help us better understand what is meant by a complete grammar for a bilingual.

Over the past ten years or so a number of researchers have compared the language abilities of heritage and adult L2 speakers at different proficiency levels in order to answer questions related to critical/sensitive period effects and the nature of bilingual interaction effects (Bruhn de Garavito, 2002; Lynch, 2003; Montrul, 2005). Recently, Montrul (2008) presented a review of such work and proposed that the variability that characterizes the development and outcomes for many early bilingual speakers should not be viewed from a deficit perspective (i.e. Schlyter’s Weaker Language as L2 proposal (1993)) because these are native speakers who happened to have a different contexts of exposure to their first language as compared to monolinguals. Even though the genesis of her work leading to the Weaker Language as L1 proposal (Montrul, 2008) is different than the one described above for the current study, and that she continues to characterize the monolingual-divergent outcomes exhibited by heritage and L2 speakers as incomplete, the general objective is the same: to get closer to an understanding of the nature of variability in early- and late-acquired grammars and to reinforce the notion that age as a linguistic
variable implies dealing with linguistic, cognitive, contextual, social, and emotional (among other) realities that can limit or enhance language learning in childhood and adulthood.

1.3 Empirical Scope and Organization of Thesis

As outlined in the previous section, this study aims to examine two aspects of monolingual and bilingual knowledge of clitic-climbing in Spanish: 1) categorical constraints of Spanish clitic-climbing; and 2) non-categorical constraints of Spanish clitic-climbing. In the latter case, a relatively high degree of variability is expected for all groups of speakers given that not one factor being examined (lexical or semantic) is expected to categorically limit the preference or use of pre-verbal clitics. Nevertheless, in comparing monolingual and bilingual speakers, the aim is to examine the distribution of this variability in different groups of speakers in order to understand how the nature (monolingual vs. bilingual) or age of exposure (native vs. non-native) to language contributes to the ways in which these speakers have acquired and treat grammatical optionality as fluent adults.

First, in the case of the knowledge that speakers have of categorical constraints that limit Spanish clitic-climbing, two questions are addressed in this study:

1) Do early and late bilinguals of Spanish permit clitic-climbing with a similar subset of infinitive-selecting verbs as monolinguals?

2) Do early and late bilingual speakers of Spanish know that embedded negation acts as a categorical constraint blocking the Spanish-specific structural option (verb incorporation, proclisis) available for clitic-climbing constructions?

Second, in the case of the variability that Spanish speakers exhibit for this optional grammatical domain, the following questions are addressed:

1) Do all groups (monolinguals, early bilinguals, late bilinguals) choose proclisis (verb incorporation) at similar rates in production, and if not, who applies knowledge of this more marked grammatical option most in production?
2) Are early and late bilinguals sensitive to the type of matrix verb (lexical factor) in limiting their use of the Spanish-specific structure (verb incorporation, proclisis) in production?

3) Are early and late bilinguals sensitive to the type of matrix verb (lexical factor) in limiting the availability of proclisis in the same way as monolinguals when analyzing the grammatical optionality?

4) Do monolingual and bilingual speakers exhibit a similar type of task-based asymmetry where their knowledge of lexical constraints of the optionality is concerned?

5) Are monolingual, early and late bilinguals sensitive to interpretational cues (semantic factor) in limiting the availability of proclisis when analyzing the grammatical optionality?

The above questions relate to the ways in which different groups of speakers have acquired and communicate using clitic-climbing structures in Spanish. This will contribute to a consideration of the following issues that fall under linguistic, psycholinguistic and sociolinguistic frames of inquiry:

1) Do speakers of a late-acquired L2 have access to the same linguistic information/mechanisms (lexical processes, computational system, UG, etc.) of the language system as native speakers of a language?

2) Are heritage and L2 speakers of a language able to acquire categorical constraints of a grammatical domain for which those constraints are not strongly represented in input?

3) Are heritage and L2 speakers of a language able to establish non-categorical constraints on grammatical optionality through probabilistic information in input related to word order variation?

4) To what degree does the bilingual speakers’ stronger experience with the unmarked structural option appear to affect their organization and limitation of the optionality in Spanish, as compared to monolinguals?
5) What age-related factors may contribute to any variability seen in the ways that monolinguals and bilinguals organize the grammar and choose variants for an optional grammatical domain?

The study is organized as follows. In Chapter 2, I present the generative and performance-related models to be considered in this study: Minimalism (Chomsky, 1995), Levelt’s (1989) production model and proposals regarding grammatical variation is represented in a language system. Here, I will outline a language model with the objective of incorporating all of these components into one language system. Further to this more broad theoretical consideration, Chapter 2 concludes with a morphosyntactic consideration of English and Spanish object pronouns and infinitival contructions in order to determine what Spanish speakers have to learn to acquire clitic-climbing constructions and what bilinguals speakers have to know in order to differentiate between grammars.

In Chapter 3, I provide a review of bilingual language effects and the language-internal and language-external factors that have been shown to contribute to non-monolingual-like linguistic behaviour. Further, I review the psycholinguistic work that has aimed to explain the cognitive underpinnings of language selectivity in early and late bilingualism, and I adapt this information to the language model outlined in Chapter 2 to provide a systematic account of the different types of language behaviour common to child and adult bilinguals (in development and at mature stages).

In Chapter 4, I review issues related to the development and attainment of language knowledge by early and late bilingual speakers. Specifically, I consider age of exposure to bilingualism as a macro-variable according to the different cognitive and contextual factors associated with language development and use as a child or adult: differences related to cognition and differences related to how children and adults are most commonly exposed to language. Further, I review empirical work that has considered the development and attainment of language in early and late bilingualism, and more specifically the acquisition and use of pronominal clitics with simple and complex predicates by speakers of Spanish in order to construct hypotheses for the current study.
Chapter 5 presents the methodology of the current study, along with the hypotheses for the research questions outlined above. Chapter 6 presents the results from the testing of monolingual, heritage and L2 speakers of Spanish, and Chapter 7 provides a discussion of those results focusing on how different types of age-related factors may be able to account for the ways in which these different groups of Spanish speakers deal with this type of grammatical optionality.
2 Language Theory

2.1 Introduction

This chapter is organized into two sections which aim to present the following:

1) To provide a theoretical background of the cognitive mechanisms that underlie language knowledge and use and how language variation can be accounted for in a linguistic system;
2) To review theoretical proposals of the morphosyntactic domains relevant to clitic-climbing structures: object pronouns and complex infinitival predicates.

In the first case, section 2.2 is divided into the following areas of discussion: i) language and Minimalism; ii) a performance model of language; iii) theories of language variation; and iv) a model for language representation and use. Subsequently, section 2.3 presents a review of the following: i) variation and optionality in the Spanish pronominal system; ii) theories of pronouns and Romance clitics; iii) theories of complex infinitival predicates and clitic-climbing; and iv) the learning task for early and late Spanish-English bilinguals.

2.2 Language and Cognition

2.2.1 Language and Minimalism

As outlined in Chapter 1, I adopt one of the basic assumptions of a generative approach to language for this study: that there is some language-specific cognitive component (language faculty) that is responsible for the development and representation of language knowledge (sounds, forms, structure). According to its proponents, this type of approach can best account for the fact that normally-developing humans acquire language in a relatively uniform way and develop language knowledge that exceeds the input that is available to them. Consequently, Universal Grammar (UG) will be the basic theory that accounts for the initial state (S°) of all normally-developing humans as they are born and begin to take in and analyze primary linguistic data (PLD) to set the constraints of the particular language that they will come to know and use (Chomsky, 1995). This linguistic component (regulating a speaker’s competence) is part of a larger performance system that uses the language knowledge of the speaker to comprehend and produce linguistic expressions. In this section, I consider the make-up of the linguistic component
and how, under a Minimalist approach, language knowledge and development (L1 and L2) can be theorized in the most simple and efficient way given the evidence that we have to date.

Here, the basic generative structure of the linguistic component of a larger language system is adopted: i) a lexicon, which is a language-specific inventory of the lexical and functional elements (and associated properties) that make up a speaker’s language; and ii) a computational system, which serves to associate the lexical and functional elements in a form that is interpretable to both the systems of articulation and interpretation of the speaker. These systems, known conventionally as PF (phonetic form) and LF (logical form), traditionally serve as the way in which the linguistic component of the system interfaces with the sensorimotor (perception and articulation) and conceptual (meaning) components involved in language comprehension and production. Language knowledge, thus, arises in being able to identify which options available in UG are most compatible with the evidence a speaker has for those sounds, forms or structures in input, and specifying those in the lexicon. Here, given that the consideration of phonological knowledge is outside the scope of investigation, it will not be considered how phonological knowledge is represented or develops.

In terms of morphosyntactic knowledge, there have been a variety of proposals in generative theory related to how cross-linguistic differences are represented in this linguistic component (different grammars; different morphological specifications, etc.). Here, the most simple and efficient proposal based in Minimalism will be adopted: that cross-language variation is a function of the lexicon and that structural variation (the idiosyncratic morphosyntactic structures of a particular language) can be accounted for in the different ways in which substantive categories (N, V, Adj, etc.) are associated in a language-specific way through the specification of functional features. As such, the computational system can be simple and invariable; it will analyze the elements selected from the lexicon as long as those elements can be computed according to universal constraints (specified in UG), and as long as they are interpretable for the particular speaker at the two interface levels, PF and LF. In this way,

---

5 By an invariable computational system, I am referring to the idea that the computational component does not vary by language, but is part of the biological endowment of a nativist linguistic system. Cross-language differences arise as the result of the functional specifications to lexical items that are analyzed by the computational system and arranged in a language-specific structure for further linguistic processing.
acquisition implies establishing a language-specific lexicon that contains a restricted set of lexical and functional elements (and their associated properties) that are drawn from the speaker’s $S^o$ (initial state—UG) on the basis of the speaker’s experience with the primary linguistic data of the particular language to which they’re exposed.

A variety of questions, therefore, emerge in consideration of such a system for bilingual populations:

1) What does this linguistic component look like for early (simultaneous) bilinguals?
2) What is the initial state for late adult learners of a language?
3) Can late learners of a language ever acquire the same type of language component instantiated by early bilinguals?

In general, there is quite a lot of evidence to suggest that simultaneous bilinguals develop independent (but selectively interacting) grammars for each language to which they’re exposed prior to 2-3 years of age (Paradis & Genesee, 1996; see Chapter 4 for review of research). In Minimalist terms, that means that two sets of lexical and functional elements may have to be acquired to make up the knowledge that a bilingual child has for each of the languages to which he or she is exposed. As mentioned, the computational system (CS) under a Minimalist approach should be invariable, and as such, there is no reason to suggest that there are different computational systems instantiated for each language if this system is able to compute a universal set of lexical and functional elements. What this system must be able to do, however, is associate these elements in such a language-specific way so as to make them interpretable to the articulatory (PF) and conceptual systems (LF) that characterizes the experience of the particular speaker. Such a system may look like the figure provided in 2.1; in Chapter 3, more details regarding a more precise configuration of this system, and how it can account for evidence of local and selective bilingual language effects, will be considered.
Figure 2.1. The Language Component of a Simultaneous Bilingual

For late language-learning, one major difference characterizes the experience that early simultaneous bilinguals and adult sequential bilinguals have for the development and representation of language knowledge: the initial state. For simultaneous bilinguals, both grammars develop on the basis of linguistic information provided by UG, but for late sequential bilinguals, they start out the process of acquisition with a lexicon already specified for their L1. As such, second-language learners (children and adults) often analyze incoming linguistic data from a new language utilizing the features of the L1 lexicon (Paradis, 2004; Schwartz & Sprouse, 1996). Quickly, however, a late language learner comes to realize that, apart from the different arbitrary sound-meaning pairings of lexical categories (nouns, verbs, etc.), he or she can not analyze the incoming data because many of the functional properties of the L2 are different from their L1, and as such, the speaker has to build a new inventory of lexical and functional elements (and their associated properties) to understand and produce the new language. A main issue among those studying acquisition from a nativist perspective, thus, has been to determine the basis upon which adult speakers build this new lexicon: through L1 features only (Bley-Vroman, 1989; Schachter, 1990), or through a combination of L1 features and UG (Schwartz & Sprouse 1996; White, 2003; Montrul, 2008). The two diverging views predict two different outcomes for late L2 speakers of a language: 1) that, in principle, late learners cannot (will not) arrive at the
same endstate as children (Bley-Vroman, 1989, among others); and 2) that, in principle, late learners can achieve the same endstate as children (Schwartz & Sprouse, 1996, among others). In general, most evidence points to the use of selective L1 properties in the analysis of a late-acquired L2, especially at the earliest stages of acquisition, but eventually also target-like knowledge of many language-specific features not instantiated in the speaker’s L1 (Liceras, 1985 for pronominal clitics; see Chapter 4 for a review). As such, I assume here that adults utilize their L1 lexicon as the initial state in the acquisition of a second language. Wherever structures are incompatible with the features assigned in the L1, the speaker has access to UG in order to establish an L2 lexicon in the most language-specific way possible (i.e. Full Transfer/Full Access hypothesis, Schwartz & Sprouse, 1996). In principle, therefore, the third question outlined above—whether L2 speakers can acquire the same kind of linguistic component instantiated by early bilinguals—can be answered in the affirmative. Nevertheless, the general consensus in the field of L2 acquisition is that adult learners fail to converge to target norms for selective grammatical domains at the most advanced stages of L2 proficiency (Coppieers, 1987; White, 2003; Long, 2007). Here, recall that target norms are typically identified as those structures that one would expect a native, monolingual speaker of a language to exhibit knowledge of. However, as mentioned previously, an L2 speaker is not a monolingual at the most advanced stage of proficiency, but a bilingual. As such, the variability that L2 speakers exhibit at the so-called endstate may not be reflective of a difference in the way that native (2L1) and non-native (adult L2) bilingual speakers instantiate the linguistic component of their overall language system or have access to a universal set of language features, but rather the unique ways in which this linguistic component interacts with other language-specific and generalized cognitive systems in early and late bilingualism, as well as the unique ways in which bilingual speakers have access to primary linguistic data at different ages. These issues will be considered in Chapters 3 and 4, but for now, I will assume that, late learners of a language can access the same linguistic information as early bilinguals, and in principle, eventually instantiate the same basic configuration of the linguistic component of the language system, illustrated in Figure 2.1.

Since the lexicon is the locus of cross-linguistic differences, we must consider it in more detail. As noted in Chomsky (1995), there are two possibilities of the way that the lexicon can function in the association of lexical and functional information: 1) that the lexicon simply stores this information, and it is only associated by morphosyntactic rules once it arrives at the
computational system; or 2) that some internal process to the lexicon can establish basic relations between associated lexical and functional items (i.e. root ‘walk’ with inflection ‘ed’ for ‘walked’ in Chomsky, 1995, p. 20) which are passed onto the computational system, where they are checked, licensed and further relations are established with other lexical and functional items so the linguistic expression can be interpretable at the interface levels. In general, the notion of a more active lexicon is plausible given the degree of language automatization (procedural knowledge underlying language processing) required to convert an idea to an utterance (or vice-versa), especially for speakers of more morphologically-complex languages (see discussion in Levelt, 1989). In this way, the most highly fluent speakers, especially of morphologically-rich languages, may have lexical processes at their disposal to establish the more basic relations among functionally-related items in order to facilitate the most efficient and automatic analysis and formulation of linguistic expressions. Questions regarding configurational and operational aspects of the lexicon will be addressed in more detail below in section 2.2.2 and in Chapter 3, when I address systematic issues related to the evidence of language autonomy and interdependence in the linguistic behaviour of bilingual speakers.

2.2.2 Language Performance

Analyzing language behaviour in bilinguals involves teasing apart the operational components of the larger language system to determine what the comprehension, production and language judgments offered by bilingual speakers in mostly experimental settings can tell us about the ways that different types of bilinguals make use (or don’t make use) of the language-specific and more performance-related cognitive components that underlie language knowledge and use. In this section, I will consider how the language knowledge established in the linguistic component of the overall system is processed for comprehension or production, and how such a system can incorporate a generative-based linguistic component, such as the one described above in section 2.2.1.

Chomsky (1995) describes the language faculty in terms of (at least) two systematic components: “a cognitive system that stores information, and performance systems that access that information and use it in various ways” (p. 2). While a generative approach to language knowledge concentrates primarily on the former of the two systems, Chomsky’s characterization of what constitutes the language faculty of a speaker highlights the importance of considering the
performance-related component if one is to characterize language from a use-based perspective (typical of the field of bilingualism) as well as a knowledge-based perspective. Performance-related processes, as we will see in Chapter 3, are especially important for the examination of bilingual language behaviour because many of the unique linguistic tendencies of bilingual speakers are characteristic of their speech production.

In general, there are four stages in the processing of language for comprehension or production (see Carroll, 2008): 1) conceptualization; 2) formulation of the linguistic plan; 3) articulation; and 4) self-monitoring. Here, I am particularly interested in the second stage as this includes the processes of lexical activation and selection that lead to the construction of a linguistic expression, and as such, is crucial for our understanding of many of the bilingual language behaviours I will review in Chapter 3. Nevertheless, the stage of conceptualization is also important for the consideration of how both linguistic and contextual influences may affect lexical and structural choices made by bilinguals for an optional grammatical domain. As such, two types of language resources will be considered as they are available to speakers for the processing of language for comprehension and production: more automatic linguistic processes, otherwise known as procedural knowledge, and more controlled linguistic processes, or declarative knowledge (Krashen, 1982; Levelt, 1989, among others). As noted in Levelt (1989), the processes of formulation and articulation are largely procedural processes, while conceptualization and self-monitoring are processes affected more by declarative knowledge (i.e. short and long-term memory stores, attentional control, etc.). However, this is a general characterization and more details will be provided about the automatic or controlled nature of certain processes as the description of each component of the system is provided.

Here, the performance model of Levelt (1989) will be used and adapted for bilingual language processing for three main reasons. First, this model is compatible with Minimalism in that they both employ a lexicalist approach to language knowledge and use (i.e. use of a mental lexicon). Second, this model adopts a serial processing and an activation spreading approach to the processing of language for comprehension or production, aspects of a performance system that may provide an efficient account of many of the lexical and structural behaviours of monolingual and bilingual speakers. Finally, this model assumes an overall top-down approach to the processing of language for comprehension or production in adult speakers, an important
feature in a bilingual language system where the discourse context and other factors such as frequency effects and the congruence of language typology may contribute to bilingual language effects. Here, these factors, as they contribute to the different types of bilingual language effects, will be considered in Chapter 3; below, I will present the aspects of Levelt’s (1989) language system that account for the processing of a linguistic utterance from concept to grammatical form (or vice-versa for comprehension) and how a Minimalist grammar might function within this larger language system.

For the case of efficiency here, this system will be described in the direction of speech production: from establishing a concept to the building of structure. In the first stages of production, a pre-verbal plan is established in the conceptualization stage that includes the necessary semantic information to be linguistically expressed in the speech act. This information will include the referents involved in whatever event is occurring, along with any propositional and situational information needed to elaborate an appropriate linguistic utterance for that event. In this way, formal semantic specifications are established (thematic roles, illocutionary information, etc.) as well as situational information (speech register, pragmatic/contextual information, etc.). As mentioned above, Levelt considers much of the information accessed for the conceptualization of a preverbal message as declarative; however, Levelt also acknowledges that the formal semantic specifications of an utterance are stored as procedural knowledge and applied automatically. On the other hand, the types of semantically-related situational variables that are relevant to a particular speech act are declarative in nature; the speaker has more control over these as compared to the formal semantic properties of the intended message. These situational variables are important because they can have an impact on later choices made in the activation and selection of linguistic features to convey the intended message in a contextually appropriate way. In most generative-based language studies, these more sociolinguistically-relevant variables are outside the scope of inquiry; however, when examining the knowledge and

---

6 Here, Levelt acknowledges that language acquisition implies establishing a language-specific “procedural knowledge base” of semantic properties that need to be encoded linguistically (established through the child’s exposure to input, and so, rather, linguistic items that need to be semantically interpreted); once language acquisition is complete, the preverbal message that is passed from the conceptualizer to the formulator automatically includes all the concepts that need to be expressed linguistically for that particular language—any other linguistic decisions made according to situational variables that can affect the structure or form of an utterance are made on the basis of declarative knowledge (p. 105-106).
use of language by bilinguals, this situational information is important to consider given the range of contextual variables that fluent bilinguals often take into account (or neglect) when communicating with other speakers, and as such, is important in the consideration of what complete acquisition means when comparing different speech groups.

Once a preverbal message has been established, the output of conceptualization is passed along to be encoded grammatically and then phonologically (mediated by the mental lexicon) so that the message can eventually be articulated. Here, Levelt considers the formulation of the intended message to be language-specific: that the mechanism involved in the grammatical and phonological encoding of the preverbal message (formulator) is different depending on the language (p. 103). The mental lexicon, therefore, is a storehouse of declarative knowledge regarding the words and their properties available for analysis in the language-specific formulator. In the adaptation of Levelt’s (1989) speech model for a bilingual population, de Bot (1992) also considers different formulators to be available for each language. Here, I will diverge from this view and suggest that the mechanism involved in linguistic encoding is not language-specific, but rather a standard mechanism available to all speakers able to analyze a universal range of linguistic computations possible in natural language (UG for grammar and a universal feature hierarchy for phonology). As such, Levelt’s (1989) grammatical encoding system (part of the formulator) is most closely associated with the computational system (CS) outlined above for Minimalism. In fact, even though Levelt suggests that each language has its own formulator, he goes on to say that the grammatical and phonological encoding of an utterance are lexically-driven; that is to say, that the particular structure and form that an utterance ends up taking is related to the morphosyntactic and phonological specifications made to the lexical entries that feed the formulator. As such, by suggesting that both the formulator and the morphosyntactic/phonological specifications of lexical items are language-specific, Levelt introduces unnecessary redundancy into the linguistic component of the system. This study, therefore, will adopt the position that Levelt’s lexical entries are specified for all the necessary linguistic information (semantic, syntactic, morphological and phonological) needed for the computational system to organize that information into an interpretable utterance from a language-specific perspective, making the linguistic component of this system compatible with a Minimalist-type grammar. Cross-language differences, therefore, are specified as features on the lexical/functional items themselves, and arise according to two factors: the specification of
semantic properties to be encoded linguistically in the procedural knowledge base of the
conceptualizer (i.e. tense, aspect, mood, deixis, etc.), and the language-specific nature of that
linguistic encoding (i.e. lexically, morphologically, phonologically, etc.). Consequently, the
lexicon plays a very prominent role in the language system adopted here. As such, I will further
consider two aspects of this component: the way that lexical items come to be selected for
structure-building (grammatical encoding) and other operational aspects of the lexicon.

In Levelt’s (1989) language model, the preverbal message established in the
conceptualization phase passes to the linguistic component and activates lexical entries contained
in the lexicon that are most compatible with the semantic specifications of that message. For
Levelt, these lexical entries have an internal structure made up of semantic and syntactic
information (the lemma) and morphophonological information. Further, one entry can include a
list of morphologically-related forms that are available for that particular item (i.e. all the
morphological variants of the verb ‘to give’, p. 191). Here, I adopt a more generative-based view
of the lexicon: that the entries contained therein are bundles of features that represent either
substantive (lexemes in a morphological approach, i.e. Beard 1987) or non-substantive categories
specified for the necessary linguistic information (semantic, morphosyntactic, phonological) as
determined by a speaker’s experience with their particular language. The most efficient lexicon
is, thus, possible as only irregular lexical items need be stored individually, and a speaker is, in
principle, able to generate an infinite number of forms from a relatively restricted store of
lexical/morphological information.

For the activation and selection of these lexical entries for grammatical analysis, I return
to the model of Levelt (1989), which adopts an activation spreading approach to lexical selection.
Here, the internal organization of the lexicon is a network, where linguistically-related items
( semantically, syntactically, phonologically) have a more dense relationship of interconnection.
As such, when a preverbal message arrives from the conceptualization phase, items that match
the semantic specification of this message are activated, and this activation can spread to other
closely-related items (by form or meaning). Further, residual activation effects may apply for
some morphosyntactic or phonological forms based on a variety of factors, such as frequency
effects, morphosyntactic complexity, and the context of communicative interaction (see Carroll
2008). The selection of particular items for inclusion in the linguistic expression occurs,
therefore, for those items with the highest level of activation. Empirical support for an activation spreading approach in monolinguals comes from a number of psycholinguistic studies using lexical identification and lexical decision tasks (Swinney, 1979; Taft, 1984), as well as lexical and structural priming tasks (Bock, 1986; Fox Tree and Meijer, 1999). Evidence for this type of internal organization of the lexicon for bilingual speakers is provided in Chapter 3 when I present a more detailed description of the configuration and operation of the bilingual lexicon.

Finally, while Levelt (1989) characterizes the mental lexicon as a “passive store of declarative knowledge” (p. 185), he goes on to suggest that languages like Turkish, which involve a great deal of morphological structure, may be subject to lexical encoding procedures before the lexical information is passed on to the formulator for further grammatical encoding. Here, Levelt alludes to the fact that languages range in the degree to which morphological composition of a word is relevant to the grammatical features specified in the lexicon, and subsequently how those words are encoded grammatically. For example, he suggests that for morphologically-poor languages, like English, “only a lexical entry’s meaning and syntax are relevant for grammatical encoding” while morphology is most relevant for phonological encoding (p. 187). However, for an agglutinative language like Turkish, there must be some automatic process ongoing in the lexicon that establishes basic relations among related morphological units to facilitate the degree of encoding that must take place in such a language (p. 187). Here, we are reminded of the two ways that Chomsky (1995) suggested that relations can be established among items selected in the lexicon: 1) that they are selected and related by grammatical rules only once they arrive at the computational system; or 2) that there is some internal process of the lexicon that associates functionally-related items before passing them onto the computational system for further association. Here, I will adopt a system whereby lexical encoding procedures are available for languages that require it; that is to say, lexical encoding may be used by speakers of different languages according to the morphological complexity of that language to facilitate efficient and automatic structure-building. Learners of a particular language, therefore, must acquire the language-specific properties associated with the lexical entries of that language; part of this functional specification will determine whether morphological information should be analyzed prior to grammatical encoding (similar to the lexical processes involved in compounding, etc.), which in turn will lead to a lexical encoding procedure that establishes more basic functional relations among related lexical/morphemic units.
This lexical encoding procedure will be addressed further in Chapter 3 in terms of the consequences this process may have for bilinguals who speak two languages that vary along the morphological complexity continuum, and for the late acquisition of a morphologically more complex language (i.e. English learners of L2 Spanish).

2.2.3 Theories of Language Variation

One further question to address in consideration of the cognitive aspects of language is how, in a simple, efficient and economical system, a speaker can represent competing (optional) grammatical forms as part of the core linguistic knowledge of the language. In most cases, grammatical optionality is conditioned by sociolinguistically-relevant factors, such as stylistically or pragmatically motivated variables. Nevertheless, for the grammatical domain under investigation here (clitic-climbing structures in Spanish), the grammatical optionality can best be characterized as “inherent variation”\(^7\) (Lavandera, 1978, p. 174); it is not only conditioned by a variety of language-internal and language-external variables, but represents a type of optionality that is available in the core grammar of Spanish regardless of the linguistic or situational context. A Spanish speaker, given the intended message, knows that there are two ways to grammatically encode this message (with pre-verbal clitics or with post-verbal clitics) without interpretational consequence. This, of course, presents a problem for a generative-based linguistic system like Minimalism where economy principles restrict this structural optionality (Chomsky, 1995; see discussion of blocking effects in Embick, 2008). Although a variety of models have been proposed to incorporate variability into syntactic theory (see review in Muysken, 2005), I will review four basic proposals here.

First, in an early view of syntactic variability, a system of rule-ordering was proposed to accommodate evidence of the English alternation in (7):

\[
\begin{align*}
(7) & \\
    & \text{i) the child whom you saw} \\
    & \text{ii) the child who you saw}
\end{align*}
\]

\(^7\) Lavandera defines “inherent variation” in the following way: “there are no speakers who never use a variant nor are there any who always use it” (p. 174). This concept could apply to allophonic variation where a speaker may use one of a number of two or more optional phonetic variants given the same linguistic context and meaning.
Klima (1964) proposed that Case marking and wh-movement were rules that could alternately be ordered by a speaker: either Case marking before wh-movement as in (7i) or wh-movement before Case marking as in (7ii). This particular model is difficult to account for under a Minimalist model of grammar because there is no clear reason why a speaker would choose one variant over the other. Here we would need to have some mechanism in place by which a speaker would have probabilities built into the system where one particular form would win out over the other for a given context of use. This mechanism would help explain frequency-related effects for such competing variants in production. Nevertheless, this early variationist proposal is difficult to systematize in a formal model of grammar because it is not clear how the structural optionality is grammatically encoded.

A second approach to accommodating grammatical optionality relies on the notion of optional rules within the syntactic component of grammar; variability results from the application (or non-application) of a particular rule. The notion of a ‘variable rule paradigm’ (Labov, 1966) was adopted to explain morphosyntactic variability, including clitic-climbing structures in Italian and Spanish (Rizzi’s (1978) restructuring rule; Aissen & Perlmutter’s (1976) clause reduction rule, respectively). Incorporating optionality into the syntactic component of grammar, however, results in a computational system that varies cross-linguistically according to the derivational options available in each language; the notion of variability (the choice made) is also embedded in this component as probabilities associated with the variants in question. This would, thus, produce a computational system that varies idiosyncratically from speaker to speaker based on the experience they have with the language in order to establish the probabilities associated with either applying the optional rule or not for any particular context. Further, there is no clear reason why any optional rule would apply if the more unmarked variant was already well-formed; this produces an inefficient system overall and introduces undesirable complexity into the computational component of grammar that is not compatible with the Minimalist framework described earlier in section 2.2.1.

To resolve the issue of complexity within the grammar itself, and remove optionality from the computational system, researchers proposed an approach in which optionality is the result of competing grammatical systems (Bickerton 1971; Kroch 1989; 1994). As noted by Embick (2008), this way of explaining competition for use (where two grammatical forms/structures can
be used for a given meaning) is most compatible with a generative system, which must respect principles related to the categorical application of linguistic rules (e.g. No Gradience principle, p. 64). While the notion of competing grammars may resolve the issue related to blocking effects in a generative system, this proposal does not clearly explain a) what multiple grammars means from a cognitive/systematic point of view; and b) how cross-language variability can efficiently be accounted for in multilingual individuals who have competition from not only competing cross-language grammars, but multiple options within any one grammar based on inherent and stylistically constrained variability. Given the range of variability that a grammar must accommodate (sociolinguistically-motivated variability, core inherent variability, bilingual variability, etc.), one speaker would have to be equipped with (much) more than ten grammars just to be able to account for all of the options available to him or her (sociolinguistically-regulated or not). This type of language system is far too redundant to consider for contexts of bilingualism (or multilingualism), and as such, will not be adopted for the current study.

The final approach to incorporating systematic variability into grammar, and the preferred model for the current study, relies on the framework outlined by the Minimalist Program (Chomsky, 1995) and the view of grammatical variability presented in Adger & Smith (2005). As outlined above, in a Minimalist grammar morphosyntactic variability is determined by properties specified to lexical items; these lexical items are specified for certain grammatical features which have an internal structure of their own: they are either interpretable or uninterpretable. Uninterpretable features need to be checked with a matched value and deleted during the syntactic computation (grammatical encoding) in order for the sentence to be well-formed; the checking of these uninterpretable features establish syntactic dependencies, but have no influence on semantic interpretation. Interpretable features, on the other hand, have a direct influence on semantic interpretation. Once certain lexical items have been selected and pass to the syntactic computation, three syntactic operations apply to build the structure—Merge, Agree and Move—depending on the features that are specified to the lexical items. The output of this computation (Spell-Out) becomes the subsequent input for morphological insertion, and ultimately, a language-appropriate phonological form. Both interpretable and (checked) uninterpretable features can affect the eventual phonological form, but it is only the interpretable features that can affect the eventual interpretation of the sentence; consequently, structural optionality may arise when the output of the syntactic computation is represented by the same interpretable
features (same meaning), but by different uninterpretable features (different syntactic dependencies). The optionality is, therefore, systematized, but also originates in a component of grammar (the lexicon) where both language-internal and language-external influences may affect the ultimate selection of one syntactic form over another.

Here, the notions of competence and performance are relevant to the linguistic knowledge that any one speaker has for grammatical optionality. Under the system outlined above, if two structural options existed in a grammar for the same interpretation, but had different probabilities associated with their use, then a fully-competent speaker should acknowledge that both structural options are as equally well-formed for a given context; choice appears to be more a notion of the certain probabilities attached to the particular structural variant for language use, and therefore, of a speaker’s selection process in a production-related task. Certainly, all analyses that have examined the issue from a variationist perspective are those that assess the probabilities associated with using one variant over the other according to certain language-internal and language-external factors (Davies, 1995, Myhill, 1988). Further, Adger & Smith (2005, p. 164) point out that a number of factors can affect the choice of a particular lexical item, including the ease of lexical access and the ease of processing, factors most often associated with real-time use of language (i.e. production). Analyzing grammatical optionality for any group of speakers, thus, implies assessing two aspects of their language knowledge: if the speaker has the optionality systematically represented in their grammar (language competence), and how that speaker chooses one structural option over the other according to a variety of language-internal and language-external factors (language performance).

In the current study, therefore, I will assume that the optionality is represented syntactically according to the features that are specified to the lexical items that are selected to fulfill a particular conceptual plan. In the case of the preverbal message generated for a variety transitive infinitival predicates (i.e. aspectual and modal predicates), a subset of infinitive-selecting (matrix) verbs is marked to establish a syntactic dependency with an infinitive (see 2.3.3) prior to the stage of grammatical encoding. The probability of this syntactic dependency occurring prior to grammatical encoding (i.e. association of functional-related items in lexicon) has to do with a variety of language-internal and language-external variables, including the type of matrix verb, characteristics of the object argument (animacy), processing speed/efficiency
(lexical access and selection, use of procedural or declarative knowledge in structure-building, etc.), mode of production (speech vs. writing), etc. I will present further details of this grammar in the review of theories of complex infinitival predicates below in section 2.3.3.

2.2.4 A Model of Language for Representation and Use

To summarize, this study will adopt a Minimalist model (Chomsky, 1995) for the representation and computation of linguistic knowledge (here the discussion is limited to the morphosyntactic domain). This grammatical system is composed of two parts, the lexicon and the computational system; the latter functions on the basis of automatic/procedural knowledge guided by innate linguistic principles of a universal grammar; as such, any variability that occurs in the language system is external to this component. The lexicon, therefore, is one of the primary loci of language variability; for any particular language, a speaker will use evidence from input to determine if and how formal semantic features are expressed linguistically. The internal structure of the lexicon is composed of a network which functions on the basis of an activation spreading system (i.e. Levelt, 1989). Once a semantic plan activates the lexicon, those lexical/functional items that best match the formal semantic features (procedural information) and situational information regarding discourse context (declarative information) will be selected for grammatical encoding. Depending on the morphosyntactic features (language-specific) instantiated to the lexical/functional items, the computational system builds the structure in such a way so that morphophonological process can apply to satisfy the conditions set out for PF and LF (for that particular speaker). While the lexicon is most guided by declarative knowledge (hence, the relative increasing or decreasing of certain lexical/functional competitors on the basis of situational knowledge regarding the speech act), some procedural processes are possible to establish syntactic dependencies among closely-related lexical/functional items prior to grammatical encoding. As such, a lexical encoding procedure may be required for some languages to ensure an efficient functioning of the system depending on that language’s degree of morphological complexity. Language variability can, thus, be built into the system in a number of ways. First, lexical/functional competitors may be able to grammatically satisfy a particular meaning (formal semantic features) in the same way (and thus, are activated at the same level), but declarative knowledge that passes to the lexicon may heighten the activation of one of the competitors to such a level as to be selected over the other. This may be the case when different
structural variants compete on the basis of sociolinguistically-relevant cues (formality of speech act, pragmatic effects, etc.). Second, when there are two core forms that can satisfy one interpretation, factors such as processing speed, ease of lexical access, frequency effects, priming effects, etc. may alter the relative activation levels of one variant over the other. Here, we may see idiosyncratic effects on an individual basis, where one native speaker of a language has a particular preference for one variant that diverges from another speaker of the same speech community (i.e. use of *dreamt* and *dreamed*). Third, while a system such as Levelt (1989) doesn’t support a bi-directional flow of information for processing, research by Kroll and colleagues on bilingual adults suggests that phonological information can have an effect on lexical choices made during selection (see review in Kroll et al., 2006). If this morphophonological information is available for the process of lexical selection, then this is an added factor that may contribute to the relative increase or decrease in activation potential of lexical competitors. These aspects of the language system will be considered in more detail for bilingual populations in Chapter 4.

In order to give a clearer picture of the language system outlined above, I have constructed a language model in Figure 2.2, one that adapts the linear model presented in White (2010), to illustrate the various components of this system.

---

8 Levelt (1989) proposes a serial model where processing happens in one direction only; information from a later phase of processing (in production) can’t influence earlier phases. In principle, therefore, phonological information cannot have an effect on the activation and selection of lexical items.
2.3 Pronouns and Infinitival Predicates in English and Spanish

2.3.1 Variation and Optionality in the Spanish Pronominal System

Before considering the theories regarding the morphosyntactic status of object pronouns and complex infinitival predicates in English and Spanish, I will review the basic distributional facts of Spanish object pronouns that are relevant to the theoretical discussions to follow. In general, Spanish object pronouns are weak functional elements (clitics) that have a variable distribution depending on tense factors and predicate complexity. In the first place, when a simple tensed verb selects a pronominal object argument, that pronominal argument is expressed in the pre-verbal position, as in (8); however when a simple predicate is untensed, as in (9), the clitic is expressed in the post-verbal position:

(8) Juan la escribe. / *Juan escribela.
    John it writes       John writes it
    “John writes it.”

(9) Juan fue a la biblioteca para escribirla. / *para la escribir.
    John went to the library to write it / to it write
    “John went to the library to write it.”
Secondly, when a speaker uses a complex infinitival predicate in Spanish and the infinitive selects a pronominal object, the pronominal clitic can appear in one of two positions: either in the canonical post-verbal position, right-adjacent to the infinitive that has selected it as argument (enclisis), as in (10i), or it can ‘climb’ up to the higher clause and be positioned left-adjacent to the tensed matrix verb of the complex predicate (proclisis), as in (10ii), but not left-adjacent to the infinitive verb, as in (10iii):

(10)  

i. Marta quiere leerla.  
*Marta wants to read it* 
“Marta wants to read it.”

ii. Marta la quiere leer.  
*Marta it wants to read* 
“Marta wants to read it.”

iii. *Marta quiere la leer.  
*Marta wants it to read* 

iv. Marta evitó leerla.  
*Marta avoided to read* 
“Marta avoided reading it.”

v. */??Marta la evitó leer.  
*Marta it avoided to read*  
[example in (v) from Strozer. 1976]

Not all matrix verbs, such as the example in (10v), support the language-specific variant (proclisis) exemplified in (10ii); the inventory of matrix verbs that do permit this variant is generally idiosyncratic to each speaker. Nevertheless, as many researchers of this domain have pointed out, the group of verbs that do permit this clitic-climbing is pretty standard across the languages that allow it (i.e. Spanish, Italian): they tend to be matrix verbs with auxiliary-, aspectual- or modal-like properties (Strozer, 1976; Davies, 1995).

---

9 While I will not consider it in any detail here, complex gerundial predicates function in the same way as complex infinitival predicates: the pronominal object selected by the untensed gerund can appear in either the canonical post-verbal position, as in (a), or in the language-specific pre-verbal position, as in (b). For study on the treatment of this variation with gerundial predicates, see Cacoullos (1999).

(a) Juan está escribiéndola. 
(b) Juan la está escribiendo. 
“John is writing it.”
Although this optionality is a core grammatical property of any Spanish dialect,\textsuperscript{10} the distribution of the clitic is not random. Davies (1995) summarizes previous research on clitic-climbing and provides one of the first large-scale corpus studies of Spanish monolingual speech and writing to statistically determine which language-external and language-internal factors influence the relative proportions of proclisis and enclisis used by Spanish speakers. In the case of language-external factors, he examined the influence of regional varieties and mode of production (spontaneous speech or writing).\textsuperscript{11} For regional varieties, he found that there was no significant difference (<10\%) in the average use of proclisis (speech and writing) by Spanish speakers from ten countries,\textsuperscript{12} but for the mode of production he did find that the rates of proclisis were considerably higher in oral production than in writing (p. 372).\textsuperscript{13}

In the case of language-internal factors, Davies considered properties of the pronominal clitic and properties of the complex predicate. In the first case, he found, as previously described (Spaulding, 1927; Rosen, 1989, cf. Davies 1995; Myhill, 1988), that characteristic of the clitic (animacy, reflexivity and use of clitic clusters) had an impact on the relative proportions of proclisis used by Spanish speakers: in general, higher rates of proclisis for animate clitics and clitic clusters (vs. non-animates and single clitics), and lower rates of proclisis for reflexive clitics (vs. non-reflexive). Further, as mentioned in Chapter 1, one of the most influential factors affecting the relative proportion of proclisis was the matrix verb employed in the complex

\textsuperscript{10} Along with these cases of optionality, there are three main types of complex predicates in Spanish where the word order of pronominal clitics is restricted to the pre-verbal position (proclisis) position only: causative constructions (I \textit{made him read it “Se le hice leer”}, perceptual constructions (I \textit{hear them singing “Los escucho cantando/cantar”}) and past participle constructions (I \textit{have read it “La he leído”}). I will not consider these constructions further, except to say that they apparently present a further challenge for the learnability of the overall Spanish pronominal system.

\textsuperscript{11} In his study, Davies refers to the mode of production as “variation by register”, and as such, assumes “the uncontroversial notion that the spoken register of a language represents the more popular tendencies of a language than the conservative written register” (p. 372). Here, I will characterize this difference as a mode of production and assume that register variability can occur in the speech practices of one speaker, including more popular, informal variants (reflected more in spontaneous, casual speech) and more conservative, formal variants (reflected in more professional, corrected speech) where applicable.

\textsuperscript{12} Speech and writing corpora examined for geographical dialect in Davies (1995): Mexico, Spain, Puerto Rico, Cuba, Colombia, Argentina, Venezuela, Chile, Peru, Bolivia.

\textsuperscript{13} Davies states (italics mine), “the average for [+CC] \textit{(proclisis)} in the spoken register is nearly three times more common than in the written register” (p. 372).
predicate. Davies suggests, as other have (Rosen, 1989 cf. Davies, 1995; Napoli, 1982), that the verbs that permit the highest rates of proclisis in spontaneous speech are those with an auxiliary, aspectual or modal function, and as such, are more light verbs in terms of their lexical semantics. Nevertheless, none of these other investigators have considered the implications that this semantic-based characterization has for the use of proclisis when the same matrix verb varies in its interpretation from a light auxiliary-like function (future tense of *ir a*) to a heavy lexical function (motion interpretation of *ir a*) depending on the context of use. As such, the range of verbs will be considered here as a lexical factor that affects the relative proportion of proclisis in the production of Spanish speakers. Those verbs tend to lie across a semantic continuum from those with less lexical content to those with more lexical content in terms of the degree to which Spanish speakers tend to use the highest to lowest rates of proclisis, respectively. For example, if we follow the results of Davies, then we would find the following continuum in (11), from the highest rates of proclisis to the lowest rates of proclisis in the spontaneous speech of monolinguals (from Davies, 1995; see complete list in Appendix A):

\[
(11) \quad \text{*ir a*} \ 86\% > \text{poder} \ 60\% > \text{querer} \ 47\% > \text{necesitar} \ 33\% > \text{preferir} \ 15\%^{14}
\]

To acquire this lexical factor is, therefore, to acquire a subset of infinitive-selecting verbs that range on a continuum (in terms of the rate of use of proclisis) from a more light function (auxiliary, modal verb) to a more heavy function (with strong independent lexical semantics). As such, speakers also must come to know which infinitive-selecting verbs fall outside of this subset despite having the same morphosyntactic status (subject control and subject-raising) as those matrix verbs that do permit the optionality. For example, Strozer (1976) suggests that the verb *evitar* (to avoid) is an infinitive-selecting verb (subject control) that is not a restructuring verb; that is to say, it does not support the language-specific word order variant (proclisis) available for other subject control verbs in Spanish, such as *querer* (to want to) and *preferir* (to prefer to).

In observing the rates of proclisis from Davies’s study above in (11), a question regarding frequency effects emerges as a potential factor that may affect the relative proportion of proclisis=enclisis across different verb types. Although no study has empirically assessed this

---

14 *Ir a* (to be going to), *poder* (to be able to/can), *querer* (to want to), *necesitar* (to need to), *preferir* (to prefer to).
factor, Davies provides raw counts of the total number of times a matrix verb appears in the
corpora he is examining, along with the percentage use of that verb with proclisis. I have
summarized selected data from Davies (1995) in Table 2.1 to show that the six most frequent
verb types used by Spanish monolinguals in Davies’s speech corpus are not verbs with the
highest rates of proclisis; in turn, the verbs with the highest rates of proclisis are not very
frequently-used expressions.

Table 2.1

<table>
<thead>
<tr>
<th>Frequency count of matrix verb</th>
<th>Proclisis by verb type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb type proclisis</td>
<td>Raw count</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>1. poder</td>
<td>3,633</td>
</tr>
<tr>
<td>2. ir a</td>
<td>2,838</td>
</tr>
<tr>
<td>3. tener que</td>
<td>984</td>
</tr>
<tr>
<td>4. querer</td>
<td>798</td>
</tr>
<tr>
<td>5. deber</td>
<td>473</td>
</tr>
<tr>
<td>6. haber que</td>
<td>385</td>
</tr>
</tbody>
</table>

These observations suggest that while highly common expressions (like ir a and poder) tend to
bias the use of proclisis, frequency effects don’t appear to be operating as a prominent limiting
variable of the word order optionality in speech production.

Further, despite the traditional semantic characterization (mentioned above) of the way
that a change in matrix verb affects the overall use of proclisis in clitic-climbing languages
(Davies, 1995; Napoli, 1982, the strict notion of semantics being a limiting factor of the word
order optionality of clitic-climbing constructions remains an open question. No empirical study
has considered the question of whether proclisis is restricted when one particular matrix verb, like
ir a (to be going to) varies in its interpretation on the light to heavy semantic continuum. In

---

15 Poder (to be able to/can), ir a (to be going to), tener que (to have to/must), querer (to want to), deber (to have
to/must), haber que (to have to/must—impersonal).

16 Ir a (to be going to), acabar de (to have just ___), haber de (to have to—obl.), volver a (to ___ again), soler (to
usually ___), poder (to be able to/can).

17 Haber que, or hay que, an impersonal expression of obligation, is a verb type that does not typically fall within the
subset of infinitive-selecting verbs that permit clitic-climbing in Spanish.
Davies’s study, for example, *ir a* is the matrix verb with which Spanish monolinguals use the highest rates of proclisis (in speech and writing), but Davies does not make any comment regarding the dual functions that this matrix verb can have: as a light auxiliary verb denoting tense (future) and agreement information for the predicate, as in (12), or as a heavy lexical verb denoting movement towards another action, as in (13).

(12) **Q:** ¿Qué haces esta semana?  
What are you doing this week?  
**A:** Mañana voy a lavar mi carro.  
[future marker]  
Tomorrow I’m going to wash my car.

(13) **Q:** ¿Adónde vas ahora?  
Where are you going now?  
**A:** Voy a lavar mi carro.  
[motion verb]  
I’m going to wash my car.

In Davies’s study, there seems to be an implicit assumption that *ir a* is used mainly as an auxiliary (tense expression) denoting future events, as it lies at the extreme light end of the light-heavy continuum of matrix verbs. Since no study has examined whether the relative light-heavy semantics of one matrix verb contributes to the proportional distribution of proclisis in the production of clitic-climbing constructions, the current study will aim to provide some much-needed empirical evidence regarding the role of the semantic interpretation of the matrix verb as a limiting factor associated with the use of proclisis in Spanish.

Finally, while the lexical and semantic variables described above tend to act as non-categorical limiting influences of the word order optionality with Spanish clitic-climbing constructions (i.e. some variability always evident), there is one main constraint that operates to categorically limit the use of proclisis with matrix verbs that usually permit the optionality: the interpolation of lexical material between the matrix and infinitival verbs of the complex predicate. For example, evidence of the way that infinitival negation acts as a blocking effect of proclisis, as in (14), along with the fact that a pronominal argument can not climb out of a subordinate clause (i.e. nominal sentential complement), as in (15), has lent support to the variety of theories, presented below in section 2.3.3, that an underlying syntactic difference differentiates word order variants (proclisis and enclisis) of a clitic-climbing structure.

(14) *Juan lo quisiera no saber pro.*  
John it would like not to know.  
Target: Juan quisiera no saberlo.
If the ungrammaticality in (14) is due to a locality constraint blocking the licensing of the clitic in the higher clause, then Spanish speakers should be able to identify the ungrammaticality in (14) (intuitively) due to their knowledge of syntactic principles associated with locality requirements (licensing of a pronominal clitic). In principle, this should be the case even if these structures are not used commonly in the speech practices of Spanish speakers (i.e. not a lot of experience with these structures in input).

2.3.2 Theories of Pronouns and Romance Clitics

In acquiring a clitic-climbing grammar, Spanish-English bilinguals are responsible for learning that the categorial status of direct and indirect objects pronouns is different in English and Spanish. In general, the distributional facts of English object pronouns have led to a relatively straight-forward analysis of the pronominal morphosyntax of these elements: that they are strong, phrasal (XP) elements subject to the same grammatical restrictions as their referential counterparts. In the examples (16)-(19) below, the pronominal object is positioned in the canonical argument position, right-adjacent to the verb that has selected it as complement, despite any variation in the tense conditions or structural complexity of the predicate.

(16) i) Peter bought DP the book.
   ii) Peter bought DP it.

(17) i) Peter needs money in order to buy DP the book.
   ii) Peter needs money in order to buy DP it.

(18) i) Peter will buy DP the book (tomorrow).
   ii) Peter will buy DP it (tomorrow).

(19) i) Peter wants to buy DP the book.
   ii) Peter wants to buy DP it.

As illustrated above in section 2.3.1, the distributional facts of object pronouns in Spanish (and other Romance languages) led researchers to propose that (unmarked) pronominal objects are not strong, phrasal elements with the same syntactic status as lexical objects, but rather are weak syntactic elements that require the support of some strong phrasal host to be instantiated.
As such, theorists like Strozer (1976) and Kayne (1975) proposed that (non-emphatic) pronominal objects in Romance languages are clitics (head constituents, X₀), defined by Zagona (2002, p. 15) as “elements which are syntactically independent words or phrasal constituents, but which are phonologically dependent.” Further support for the clitic status of Romance pronominal objects (in modern Spanish) comes from the fact that they neither directly replace the lexical DP (i.e. are not subject to the same structural constraints), as in (20), nor can they stand independently as their own phrase, as in (21).

(20) i) Pedro vio a María.
Peter saw Mary
“Peter saw Mary.”

ii) *Pedro vio (a) la. / Pedro la vio.
Peter saw her / Peter her saw
“Peter saw her.”

(21) i) ¿A quién vio Pedro?
pronoun who saw Peter
“Who did Peter see?”

ii) (A) *la. / (A) ella.
her / she
“Her.”

While there is a general consensus on the syntactic status of pronominal objects as head constituents in Romance languages (Suñer, 1988; Kayne, 1989; Franco, 1993; Sportiche, 1996), there are two diverging views of how these elements are represented syntactically: movement accounts and base-generation accounts. Before considering any theoretical proposals, I review the three main pieces of evidence that have complicated the elaboration of any comprehensive theory of Romance pronominal clitics: language-specific linearization facts, clitic-climbing and clitic-doubling. In the first case, the three most cited Romance languages for theoretical purposes are French, Italian and Spanish, and exhibit language-specific distributions of pronominal clitics. In the case of French, pronominal clitics are always positioned in the pre-verbal position, left-
adjacent to the verb that has selected them as complement, as in (22). In Italian and Spanish the position of the clitic varies according to tense specifications and the structural complexity of the predicate, distributional facts described in section 2.3.1 and summarized here for Spanish in (23). 

(22) i. Jean la veut. / *Jean veut la.  
    John it wants / John wants it.  
    “John wants it.”

ii. Jean veut aller à la bibliothèque [pour la lire.] / *[pour lire la.]  
    John wants to go to the library [to it read] / [to read it]  
    “John wants to go to the library to read it.”

iii. Jean veut la lire. / *Jean veut lire la. / *Jean la veut lire.  
    John want it to read / Jean wants to read it / John it wants to read  
    “John wants to read it.”

(23) i. Juan la quiere. / *Juan quiere la.  
    John it wants / John wants it.  
    “John wants it.”

ii. Juan quiere ir a la biblioteca [para leerla.] / *para la leer.  
    John wants to go to the library [to read it] / [to it read]  
    “John wants to go to the library to read it.”

iii. Juan quiere leerla. / Juan la quiere leer. / *Juan quiere la leer.  
    John wants to read it / Jean it wants to read / John wants it to read  
    “John wants to read it.”

Second, any theory describing the Romance pronominal system must be able to account for evidence of clitic-climbing in languages like Spanish and Italian, but its restriction in a language like French, illustrated in (22iii) and (23iii) above.

Finally, such a theory must also be able to account for a wide range of language- and dialect-specific rules on clitic-doubling. In general, most Romance languages instantiate

\[\text{Italian patterns for these Spanish examples in the same way:}\]

(a) Giovanni la vuole. / *Giovanni vuole la.
(b) Giovanni vuole andare a la biblioteca per lirla / *per la lire.
(c) Giovanni vuole lirla. / Giovanni la vuole lire. / *Giovanni vuole la lire.
restrictions on the utterance of phrasal object pronouns (resulting in an emphatic/clarifying expression), as in (24), where clitic-doubling is generally required:

(24) *(La) vi a ella.
    cl I saw prt her.
    “I saw her/HER.”

Further, in Spanish when dative arguments do not have the thematic role of recipient, the doubling of the lexical argument is required (Cuervo, 2003), as in (25ii):

(25) i. Juan compró las flores [PP para María]. (beneficiary)
     John buy-pst. the flowers [PP for Mary]
     “John bought the flowers for Mary.”

ii. Juan *[le] compró las flores a María. (beneficiary)
    John [ DAT bought the flowers prt Mary]
    “John bought the flowers for Mary.”

iii. Juan envió las flores [PP a María] / Juan le envió las flores a María
     John sent the flowers [PP to Mary] / John DAT sent the flower prt Mary
     “John sent the flowers to Mary.” [Mary = recipient]

Finally, in a variety of Spanish dialects, the doubling of the direct object argument is common, often subject to restrictions related to the semantic features of the accusative argument (i.e. animacy and specificity, Suñer 1988), as in (26):

(26) Lo salúde a Juan. (Sánchez, 2003, p. 48)
    clI greeted to Johni
    “I greeted John.”

In considering the theories proposed to account for the syntax of Romance pronominal clitics, therefore, these three distributional facts will be kept in mind: cross-language linearization facts, clitic-climbing and clitic-doubling. Further, in order to evaluate any one theory, I will comment on what any theory might predict for acquisition contexts and offer evidence, where it is available, to support or refute any of the proposals under consideration.

2.3.2.1 Movement Approaches to Romance Clitics

As mentioned previously, theories of Romance clitics are divided into two basic approaches: movement approaches and base-generation approaches. In the first case, the general view is that pronominal clitics in Romance are generated in the canonical argument position of the argument-selecting verb (as an object complement), and then move to their surface position via syntactic
principles and language-specific rules. Kayne (1975; 1989; 1991) suggests that the basic relationship between an argument-selecting verb and its pronominal clitic arises through obligatory head movement of the pronominal clitic to an invariable left-adjoining position to the verb (cl-V): as syntactic heads (X₀), clitics must stand in a left-adjoining relationship with another head (p. 647). As such, the linearization facts for French and for unmarked simple tensed predicates in Spanish and Italian can easily be accounted for. For more complex predicates, Kayne goes on to suggest that the null-subject status of a language (strength of infinitival I₀) determines a basic difference between languages like French, which don’t instantiate clitic-climbing, and Spanish and Italian, which do have clitic-climbing. As such, Kayne implies that the word order facts of French are more basic from a derivational point of view and that for all types of clauses other than simple tensed predicates in Spanish and Italian, the linearization facts arise from the application of at least one, but possibly more movement operations. In fact, Kayne (1991) proposes that the post-verbal alignment of pronominal clitics with simple untensed predicates (infinitives, gerunds) in Spanish and Italian arises from a series of movement operations: head movement of clitic to left-adjoin infinitive (basic configuration) and then movement of infinitive above the clitic (in I₀) to a single-bar projection (I’). Kayne’s (1989; 1991) proposals, while providing a simple and efficient system for French, introduces a considerable amount of syntactic complexity into the systems of Spanish and Italian. Further, his movement accounts have the clitic generated in the canonical argument position in order to pick up case and theta-assignment (leaving an empty category/trace). Accounting for clitic-doubling is difficult as there is no clear way that the lexical argument (to which the clitic is referring) could appear in the same clause, nor how the clitic would maintain a relationship with its co-referring lexical argument.

Uriagereka (1995) attempts to resolve the problem of clitic-doubling for a movement account of Romance clitics by proposing a basic syntactic difference between 1st/2nd person clitics and 3rd person clitics (based on proposals by Muysken, 1982 and Chomsky, 1992 cf. Uriagereka, 1995). By drawing diachronic and synchronic parallels between 3rd person clitics and definite articles, Uriagereka proposes an extension of the DP hypothesis to Spanish 3rd person clitics: clitics are heads (D₀) of determiner phrases generated as direct object arguments of a transitive verb. When the clitic takes on a pronominal function, pro is selected as the NP of the argument, but in the case of clitic-doubling contexts, the lexical referent is instantiated as the NP
of the argument. In Uriagereka’s account, like Kayne’s, the unmarked configuration between a clitic and its selecting verb is pre-verbal (motivated through a basic required derivation so that the clitic can receive reference in $F^o$ (between $C$ and $I$)). Again, similar to Kayne’s proposal, Uriagereka suggests that the post-verbal position of clitics in untensed simple predicates in Spanish and Italian results from an additional movement operation of the infinitive to a position above the clitic in $F^o$.

Given that the movement analyses of pronominal clitics above propose that the basic unmarked position for Romance clitics is pre-verbal (in the functional domain of its argument-selecting verb), one might expect there to be some early developmental data from the acquisition of Spanish or Italian where learners pass through a stage of using the more basic or unmarked configuration of clitics (cl-infinitive) where it is not permitted. Although not a lot of extensive analysis has been done to examine the early acquisition of clitics with complex predicates in Spanish or Italian children (<3;0), the empirical evidence available to date suggests that once children start to use clitics with simple untensed predicate and complex predicates, they do not make linearization errors. As such, there is no empirical evidence to support the notion that the configuration [cl-infinitive] is the unmarked first step of a complex series of syntactic operations required to realize post-verbal clitics in untensed simple predicates and clitic-climbing contexts. Finally, both accounts imply either complex pronominal systems (as in Uriagereka) or a complex system of derivations that vary cross-linguistically (Kayne and Uriagereka). Under a Minimalist account of cross-linguistic differences, it is difficult to adapt either of these approaches because they introduce a certain amount of language-specific syntactic complexity.

### 2.3.2.2 Base-Generation Approaches to Romance Clitics

Base-generation approaches attempt to resolve some of the complex derivational issues associated with movement approaches of Romance clitics. In general, for base-generation approaches, the clitic surfaces as an affix or, in more recent proposals, as a morphological representation of features associated to the object in a position relevant to the checking and licensing of these features. As such, language-specific linearization facts can arise from post-syntactic processes where the morphological characteristics of the verb (tensed, untensed) may influence whether the clitic surfaces in a pre- or post-verbal position. Strozer (1976) presents one of the earliest base-generation proposals and argued for a system of (fixed order) PRON(ominal)
nodes where the clitics could be generated directly. Clitics would form morphosyntactic constituents with the verb and be co-indexed to the lexical noun that could select a [+PRO] feature to instantiate the clitic.

Subsequent base generation accounts captured the basic premise of this system where directly generated clitics form a constituent with the verb and are co-indexed to a (mostly null) lexical argument, and in turn realize certain features instantiated for the object complement. For example, Suñer (1988) proposed that clitics are instantiated as agreement morphemes of the argument-selecting verb (formalized in a Matching Principle). In this account, the lexical (null) argument receives both case and theta-assignment, while the clitic “enter[s] into inflectional relations and [is] generated as part of the verb” (p. 393). For a clitic to become lexically-realized, it must be co-indexed to the lexical argument and the clitic must c-command the argument. Here, clitic-doubling can be accounted for where language- or dialect-specific rules apply, and both the clitic and lexical argument become overtly realized. However, given that Suñer’s approach has the clitic represented as “part of the verbal head” (p. 403) of the argument-selecting verb, then clitic-climbing contexts are difficult to account for. To resolve this issue, later accounts (Franco, 1993; Sportiche, 1996) proposed that clitics maintain their independent syntactic status from the verb by being instantiated as the head of some functional projection, such as AgrO. The type of agreement that happens between the object clitic and the object argument is analogous to the agreement relation that occurs between the morphological agreement on a verb and the subject (projected in AgrS). In these accounts, the null object argument has the formal status of a null pronominal object, such as pro, following Jaeggli (1986, cf. Franco 1993).

For the current study, therefore, I will adopt the approach to pronominal clitics outlined in Franco (1993), an approach designed on the basis of evidence from Spanish. In this analysis, Franco attempts to account for a universal range of cross-language differences relevant to the use of pronominal clitics by suggesting that cross-linguistically, pronominal clitics lie on a continuum between inflectional affixes and strong pronouns (i.e. phrasal elements), as defined by a variety of properties, such as adjacency constraints and constraints on fixed ordering (see table of properties in Franco, 1993, p. 67). Under this approach, the forms and linearization of pronominal clitics are not the result of inserting lexical items into syntax and moving them around to achieve the appropriate syntactic conditions for checking and licensing, but rather the
result of linguistically expressing a bundle of morphological features attached to a certain position in the inflectional field of an argument-selecting verb. The actual expression of the clitic is not related to its syntactic position, but rather to post-syntactic (PF) rules associated with the morphophonological status of the clitic (as a weak phonological element) and language-specific rules related to how such a clitic becomes aligned to its supporting verbal host. This type of systematic approach to clitics is more compatible with a Minimalist framework where the syntax is invariable, and where cross-languages differences are predominantly morphological.

Overall, base-generation accounts of pronominal clitics do what movement accounts cannot: remove cross-language variation from syntax and restrict it to the lexicon and post-syntactic processes. Further, evidence of early target-like distributions of pronominal clitics with few (if any) qualitative linearization errors in child acquisition (Domínguez, 2003; Thomas, 2007) point to a system where the speaker is not computing a number of basic to more complex steps in order to acquire a language-specific grammar, but rather has to acquire the language-specific functional features associated with the morphosyntactic status of pronominal clitics in their language (i.e. on a continuum between inflectional affixes and phrasal constituents). The unresolved question regarding the optional word order variants available for clitic-climbing structures in Spanish and Italian, however, has to be addressed since in a basic base-generation system, one would expect only one grammatical position to be available for complex predicates (AgrO of the infinitive). The following section 2.3.3 will consider how such optionality can be incorporated into a grammar with a Franco (1993)-type approach to pronominal clitics.

2.3.3 Theories of Complex Predicates and Clitic-Climbing

Using a base-generation approach to Spanish pronominal clitics makes it necessary to determine how two word order variants are instantiated in the core grammar of Spanish for clitic-climbing constructions. I consider the different theories that have attempted to account for the syntactic characteristics of clausal formation in languages that permit clitic-climbing. Further, here I examine how a wide range of verb types in Spanish with different syntactic characteristics (auxiliaries, modals, transitive verbs) come to form a subset of infinitive-selecting verbs that permit clitic-climbing. Finally, I compare this with the English syntax of these verbs in order to more precisely determine what has to be learned by an English speaker of L2 Spanish and the parts of grammar that may be (and are not) congruent for a Spanish-English bilingual.
According to traditional generative-based proposals, there are three ways that a tensed verb and an infinitive, with the same referent as external argument (subject), establish a syntactic relationship in Spanish and English: 1) functionally—the tensed verb is part of the functional domain of the infinitive (auxiliaries); 2) subject control—the infinitive is the verbal complement (CP or IP) of the matrix verb with PRO as the subject of the infinitive (co-indexed to subject of matrix verb); and 3) subject raising—the infinitive is the verbal complement (CP or IP) of the matrix verb and the subject of the infinitive raises to the subject position of the matrix clause. The grammars of English and Spanish are relatively congruent where the instantiation of these three categories is concerned; as such, having to learn the syntax of verbs like querer (‘to want to’ as a subject control verb) and parecer (‘to seem to’ as a subject raising verb) is pretty straight-forward for a speaker of English. Nevertheless, these two grammars diverge on two basic syntactic points. First, English modal verbs like ‘can, must, shall’ are typically grouped with auxiliaries in terms of their functional properties, and as such, tend to be instantiated within the functional domain of the infinitive of any given predicate. On the other hand, in Spanish these verbs are thought to pattern syntactically with main verbs (Zagona, 2002, p. 42), and as such, many analyses have a verb like poder (to be able to/can) being included in the category of subject raising verbs (Strozer, 1976). The consequences that such a syntactic difference may have for the acquisition of Spanish as a second language or for the use of complex predicates by bilinguals will be considered in more detail in Chapter 4. Second, English and Spanish diverge syntactically when the referent of the external argument (subject) of the first proposition is different than that of the second proposition. In this case, Spanish speakers typically use a sentential complement, as in (28), while English speakers continue to use infinitival complements, as in (29):

(28) i. Juan quiere [que María vea la película].  
John wants that Mary see the movie

ii. *Juan quiere María ver la película.  
John wants Mary to see the movie

(29) i. John wants [Mary to see the movie].  
ii. *John wants that Mary see the movie.
To limit the discussion here to cases relevant to Spanish clitic-climbing, I will focus on complex predicate types where the subject is the same for the tensed and infinitival clauses of the complex predicate. Nevertheless, as we will see below, the data in (28) has some relevance for the theories that describe the syntactic configuration(s) of the two word order variants (proclisis and enclisis) of a clitic-climbing structure.

Continuing with our example of querer (to want to), it is common to many languages that such a transitive verb can satisfy its selectional requirements by selecting among nominal, verbal or sentential complements, depending on language-specific constraints. In Spanish, other than a noun, querer can select an infinitival complement, as in (30), or a sentential complement, as in (31) according to constraints on the identification of the subject for each proposition:

(30)  Juan quiere [ver la película].
     John wants [to see the movie]
     “John wants to see the movie.”

(31)  Juan quiere [CP que María vea la película].
     John wants [CP that Mary see the movie]
     “John wants Mary to see the movie.”

In most accounts, the sentential complement in (31) is considered to be a full clause (CP), and as such this sentence has a bi-clausal structure; in this case, when the direct object argument of the verb vea (see) is pronominalized, it must remain in this lower clause, as in (32i), and can not appear left-adjacent to the matrix verb, as in (32ii) because of locality constraints associated with the licensing of the clitic (CP acts as a barrier to this licensing).

(32)  i.  Juan quiere [CP que María la vea].
     John wants [CP that Mary it see]
     “John wants Mary to see it.”

     ii. *Juan la quiere [CP que María vea].
         *John it wants [CP that Mary see]

While we know that both the pre- and post-verbal positions are available for the pronominalized sentence in (30)—Juan la quiere ver and Juan quiere verla—when lexical material is embedded between the matrix and infinitival verbs of the complex predicate, the language-specific word order becomes unavailable. For example, in (33i), the unmarked word order (enclisis) is possible when the infinitival predicate is negated, but the language-specific word order is not, as in (33ii):

(33)  i.  Juan no quiere [ver la película].
     John does not want [to see the movie]
     “John does not want to see the movie.”

     ii. *Juan la no quiere [verla].
         *John it does not want [verla]
John would like not to know it. [that his wife deceived him.]

If this structural restriction on the availability of the language-specific word order is compared to the data in (32), then it appears as though the restriction in (33ii) is due to the fact that embedding negation forces the underlying structure of the complex predicate to have a bi-clausal configuration, and as such, the pronominalized argument of the infinitive must remain in the lower clause. Given these structural assumptions, then when the direct object argument from (30) is pronominalized, the unmarked variant (enclisis) can also be bi-clausal, as in (34i), but the language-specific variant must have an underlying clause structure that is reduced in some way to allow the clitic to “climb” up to the higher clause, as in (34ii):

This basic assumption, of an underlying structural difference for the unmarked and language-specific variants in (34), has served as the basis for most generative-based analyses of this clitic-related phenomenon. Some of the earliest theoretical characterizations of this difference as “restructuring” (Rizzi, 1978) and “clause reduction” contexts (Aissen & Perlmutter, 1976) persist descriptively in much of the theoretical and acquisition literature today (Wurmbrand, 2001; Cinque, 2004). However, the earlier proposals, being grounded in earlier versions of generative theory, have the optionality instantiated for this domain as a syntactic rule, where the basic derivation is the bi-clausal structure (enclisis) and a language-specific syntactic rule optionally applies to restructure or reduce the predicate into a monoclausal configuration (proclisis). As mentioned previously, such a system which places the optionality and cross-language variability in the syntax is undesirable in a Minimalist system where cross-language differences are morphological and the computational system is mostly invariable. Further, it is difficult to account for such an approach in a maximally economic system where differences in syntactic structure are accompanied by some interpretational difference (see discussion in Embick, 2008). Nevertheless, the idea that the post-verbal variant (enclisis) is more basic or
unmarked from a structural point of view is a valuable insight, especially for the consideration of how such optionality is handled by bilingual speakers.

In order to place the optionality outside of syntax, then decisions on predicate formation must be made prior to the computation of the entire linguistic expression. What this has implied traditionally is that verbs do not necessarily have a uniform structure in the lexicon. Strozer (1976) develops an account where verbs that select infinitival complements can either subcategorize to select full sentential infinitival complements (CP in more current approaches) or reduced VP complements. She distinguishes semi-auxiliaries, such as modal and aspectual verbs, from querer-type transitive verbs that, apart from being able to select infinitival complements, also select nominal and sentential que-complements (unlike semi-auxiliaries). She further subdivides transitive verbs into two classes: those, like querer, that select VP infinitival complements (and, in principle, CP complements), and those like evitar and temer which invariably select CP infinitival complements (and don’t permit clitic-climbing).

Strozer’s analysis gives a clear account of what constitutes a trigger verb for clause reduction in Spanish: the verbs (marked lexically) that subcategorize for VP infinitival complements constitute the subset of clitic-climbing verbs in Spanish. Here, the structure of the complex predicate can be determined before the syntactic derivation, but it remains unclear how infinitives can be instantiated in the lexicon as both reduced (VP) and full sentential (CP) structures without introducing unnecessary redundancy (or having other syntactic consequences). For this reason, other analyses of restructuring in Spanish and related languages have attempted to eliminate the optionality and present an approach where the infinitive is the main verb of the construction, and the matrix verb has an invariably functional role in the linguistic expression.

Cinque (2004) proposes such a functional account of complex infinitival predicates in which two claims are highlighted: 1) that all restructuring is functional; and 2) that restructuring occurs in the absence of transparency effects, or rather, that the syntactic configuration of a

---

21 The framework provided by Distributed Morphology (Halle & Marantz 1993) could also account for clitic linearization facts as a post-syntactic phenomenon. This approach is not adopted here because it is traditionally understood to be a non-lexicalist approach to grammar, a framework that is not compatible with the lexicalist view of acquisition and bilingualism adopted in this study.
complex predicate is the same regardless of the presence or absence of pronominal clitics. This proposal is attractive in that it simplifies the grammar and provides a one-to-one mapping of syntax to interpretation. Nevertheless, while this proposal removes optionality from predicate formation, it reintroduces the problem of how the clitic could be optionally expressed in one of two positions (enclisis or proclisis) for the same linguistic expression (no interpretational difference). If the infinitive is the main verb of the predicate with a functionally-expressed matrix verb, then a base-generation approach would predict that there is only one position available for a pronominal clitic in a complex infinitival expression: in the AgrO of that predicate. Cinque does little to explain how two pronominal word orders are possible with this particular structure. Even if he assumes that the linearization of a clitic is a morphophonological event, then the pronominal clitic would identify this clause as tensed and invariably surface in the pre-verbal position. He doesn’t give any clear account of how the post-verbal variant is possible. In this case, only a movement approach could possibly generate two different structures: either the clitic is generated in the argument position of the infinitive (enclisis) and moves up (proclisis), or the basic configuration of a pronominal clitic is pre-verbal (Kayne 1975; Uriagereka 1995) and the infinitive moves up above the clitic. What remains unclear is why any secondary movement would occur at all if the more basic sentential version satisfies all conditions of interpretability.

For a lexicalist account of clitic-climbing to work in the most simple and efficient way, we turn to an approach by Masullo (2004), who describes the optionality with a Minimalist grammar. In this account, the optionality characterizing clitic-climbing constructions is accounted for by an order of Merge operations. The post-verbal variant (enclisis) arises through an unmarked process by which the pronoun first Merges to the infinitive that has selected it and checks its selectional features (case and theta), and then that predicate Merges to the matrix verb. On the other hand, the pre-verbal variant (proclisis) arises through a language-specific process by which the Merge of the pronoun is delayed until the two verbs Merge and undergo complex predicate formation (CPF). CPF is a feature lexically-specified to the matrix verb that results in a process akin to verb incorporation: the verbal heads incorporate and pool their argument structure while the unchecked selectional features of the infinitive (case and theta) percolate to the complex verb for later checking when the pronoun merges with the complex predicate. These two processes and structures are exemplified in (35) and (36), respectively:
Although Masullo uses a movement account of clitics in his proposal, the examples above show how a Franco-type (1993) account could work as well where pro merges in the argument position and the clitics later surface as morphological agreement markers of AgrO. The linearization of the clitics is, thus, a post-syntactic phenomenon associated with language-specific morphophonological constraints: in the case of the unmarked variant in (35), the clitic is linguistically expressed in the post-verbal position because it identifies the clause as untensed; conversely, in the case of the language-specific variant in (36), the clitic surfaces in the pre-verbal position because it identifies the clause as tensed. Further, one can also see how negation can act as a blocking effect of the language-specific variant in (37). In order to negate only the infinitival predicate, the infinitive must project some functional structure to accommodate the negative marker (NegP):
What Masullo’s system has a hard time accounting for, however, is how a language-external factor, such as register variation, may affect the relative application of CPF. For example, why should a speaker use less overall proclisis when pronominalizing the sentence *Juan va a ver la película* (John is going to see the movie) in more formal, corrected speech as compared to casual, spontaneous speech?

To account for the effect of register variation on the relative proportions of proclisis used by monolingual Spanish speakers, we return to Chomsky’s (1995) insight about the lexicon, and Levelt’s (1989) proposal regarding lexical encoding. Levelt proposed that lexical encoding may be available for languages that require a procedure by which functionally-related lexical items may become associated before those items pass onto grammatical encoding. Taking Masullo’s proposal of verb incorporation, I would suggest that a lexical encoding procedure incorporates the two related verbs as long as two conditions hold: 1) that the CPF feature is activated to a critical threshold; and 2) that the speaker’s processing for this domain can occur relatively rapidly and automatically. In the first case, a threshold for activation means different things for different verbs; in some cases, just activating the CPF feature is sufficient enough for it to apply in casual speech (i.e. for verbs like *ir a, acabar de* and *poder*), while for others, other language-internal factors may have to contribute to the overall activation effect to reach the required threshold for lexical encoding to take place in casual speech (i.e. for *tratar de* and *preferir*). In the second place, a variety of factors associated with declarative processes (language-external influences) may impede the processing speed required to have lexical encoding occur even if the CPF is activated to the required threshold: more formal and corrected speech, mode of production (writing vs. speech) and, perhaps, bilingual effects.

With lexical encoding available, we are able to remove variability from the computational system. The grammar does not have to be responsible for a delay in merging a pronominal clitic, nor does it have to be responsible for the incorporation process, nor the representation of two

---

22 Davies (1995) showed that Spanish monolinguals used very low rates of proclisis with these two verbs: *tratar de* ‘to try to’ 20% and *preferir* ‘to prefer to’ 15%.
verbal heads. As in French or English, the merge of the pronominal object in Spanish occurs as a primary computation once all the elements of the proposition enter the grammar: it either merges to the infinitive alone (enclisis), or it merges to the incorporated verb (proclisis). Further, the process of verb incorporation can work much in the same way that other incorporation processes, like noun incorporation, work linguistically. For example, the way that noun-noun compounds arise in Spanish is typically through a lexical process where the noun head selects a prepositional complement:

(38) galleta [de chocolate]
    cookie of chocolate
    chocolate cookie

The grammar does not associate these lexical items as a primary operation, but rather, the compound enters syntax as one nominal phrase with *galleta* as the head of that phrase. Similarly, then, if the verbs incorporate prior to arriving at the computational system, then the grammar only has to be responsible for an unmarked structure: the merging of a verbal head to the pronominal argument. Even though the incorporated verbal elements are already structurally associated when they enter the grammar, the syntax is able to analyze its component parts in order to check features associated to tense, person, etc. on the verbal head (the tensed verb), rather than on the infinitive.

The process of lexical encoding that forms the complex predicate works much in the same way that Masullo proposes: the argument structures of the two verbs are pooled and the selectional features of the infinitive percolate for later checking. As such, the same interpretable features are available regardless of whether verb incorporation has occurred, and consequently both structures can be interpreted the same way given the same context (re: Adger & Smith 1995). As mentioned above, it is the matrix verb that checks tense and agreement because, like in other languages that employ verb incorporation (eg. Hindi), the matrix verb (the vector/explicator) serves as a way to express subtle semantic information (i.e. tense and aspect) and mark grammatical inflection (Gerdts, 1998). Finally, this account makes it easier to explain why the verbs that lie at the extreme light end of the semantic continuum tend have the highest

---

23 This type of incorporation can be seen from a morphological perspective where the compound “serves the combined syntactic function of both elements” (Gerdts, 1998, p. 84).
rates of proclisis despite not being very frequent in input. If this system works like others that employ verb incorporation, then the verbs that contribute very subtle semantic information to the linguistic expression are more likely to enter into this incorporated relationship more easily than verbs that are heavier semantically. Nevertheless, as a lexical process, this system does not rule out the possibility that speakers can use verb incorporation with expressions that are more semantically heavy, such as saber (to know how to), and aprender a (to learn how to). As mentioned previously, the inventory of verbs that a Spanish speaker selects to employ verb incorporation is idiosyncratic. Further, the tendencies of speakers to use verb incorporation (or not) may be idiosyncratic (based on frequency effects, probabilities associated with certain lexical items, etc.) on the basis of a variety of language-internal factors (i.e. type of matrix verb, reflexivity of clitic, etc.), and on the basis of the speakers experience with different registers of the target language. For monolinguals, this basic knowledge they have of verb incorporation is highly procedural (and associated with a number of processing and experiential factors), and as such, the learning process for bilinguals and non-native speakers is how to systematize all of this information into the most language-specific grammar possible.

2.3.4 The Learning Task of Spanish-English Bilinguals

In order to acquire clitic-climbing according to the description above, Spanish speakers must learn the following:

1) The clitic status of object pronouns and their linguistic properties.
2) Verb incorporation and relevant constraints.
3) The subset of matrix verbs that instantiates the CPF feature to optionally allow verb incorporation, and the base activation level of this feature for each verb.
4) The language-internal and language-external factors that can affect the application of verb incorporation as a lexical process.

In Minimalist terms, the knowledge of the first two items is provided by Universal Grammar; once a child is exposed to Spanish consistently, he or she will restrict his or her grammar to include pronominal clitics and verb incorporation (and the relevant constraints, eg. blocking effects of infinitival negation). In the case of simultaneous bilinguals, as long as the child has enough consistent input through the first couple years of life, these children should be able to acquire these grammatical features just as monolinguals would. For adult English speakers of L2
Spanish, as long as UG remains available in adulthood, there is no principled reason why they could not acquire the clitic status of object pronouns, or the verb incorporation process of complex predicates. Where monolinguals, early bilinguals and late bilinguals may diverge, however, is in the third and fourth points above, which crucially develop on the basis of experiential evidence.

In the case of establishing a subset of matrix verbs that permit verb incorporation, each speaker establishes his or her own idiosyncratic inventory of verbs that allow it and restrict it (Strozer, 1976; Moore, 1996); however, as mentioned previously, most monolingual speakers of Spanish converge upon a similar inventory of auxiliary, modal and aspectual verbs. Nevertheless, in order for any speaker to establish this inventory of clitic-climbing verbs, transparency effects, such as the position of pronominal clitics, is an important feature that speakers must attend in order to not only decide which verbs allow clitic-climbing, but also at what rate and according to what conditions (i.e. base activation levels of CPF) clitic-climbing occurs. Further, the relative use of proclisis by any one speaker will be affected by i) the knowledge of how different language-internal conditions increase or decrease the probability of verb incorporation (based on activation levels in the lexicon); and ii) language-external factors related to declarative knowledge regarding the speech act (formality, interlocutors, etc.) that can alter how automatic the process of structure-building is: the less automatic, the less verb incorporation. In this case, for English L2 speakers of Spanish, assuming that they have access to the same kind of input as native speakers, they must not only acquire the lexical encoding process (in principle, available via UG), but also be able to filter out declarative-based knowledge (i.e. explicit learning, experience with written structure, etc.) that might interfere with the procedural-based knowledge that facilitates the process of verb incorporation. Further, for all Spanish-English bilinguals, they may have an additional language-internal influence that affects the relative activation levels of CPF in Spanish: their knowledge of English complex infinitival predicates (unmarked structure-building). The goals of Chapter 3 and 4 are to review empirical evidence on the unique language behaviours that result from different types of dual-language experiences—especially those of Spanish-English bilinguals in North America—in order to predict how early and late bilingual speakers of Spanish will perform on tasks testing their knowledge and use of clitic-climbing in the current study.
2.4 Chapter Summary

In this chapter I described the theoretical approaches that are adopted in the current study. First, I outlined a language system with a Minimalist grammar (Chomsky, 1995), a modular and spreading-activation approach to language production (Levelt, 1989) and a systematic approach to grammatical variation compatible with Minimalism (Adger & Smith 2003). This framework produces a language system that can most efficiently deal with the procedural aspects of core grammatical optionality (representation of more than one grammatical option for the same meaning) as well as the performance-related aspects of using competing grammatical variants (factors that may affect making on-line choices in production).

Second, I reviewed some theoretical proposals of Spanish pronominal clitics and complex infinitival predication. For the current study, I follow the base-generation view of pronominal clitics outlined in Franco (1993): pronominal clitics are head constituents realized as object agreement morphemes of the functional category AgrO. Further, a Minimalist approach to complex infinitival predication is adopted as outlined in Masullo (2004): either 1) the verbs incorporate first combining the argument structure of the two predicates, thus instantiating only one AgrO, and as such, proclisis; or 2) the verbs do not incorporate and the pronominal argument first merges to the infinitive, thus checking the features of the untensed predicate and surfacing as enclisis. In order to remove the optionality completely from the grammar, I consider verb incorporation to occur on the basis of a lexical encoding process. Following earlier discussions (Chomsky, 1995; Levelt, 1989), I assume that verb incorporation occurs in a pre-syntactic phase where related functional material are associated morphologically; the merge of the pronominal argument to the predicate is, thus, a primary syntactic operation.

In the Chapters 3 and 4 I expand on the theoretical discussion here in three ways: 1) how the language model above needs to be adapted for contexts of dual-language exposure on the basis of evidence of language knowledge and use in bilingual contexts; 2) the factors that affect the process and outcome of acquiring typologically distinct grammatical systems in contexts of bilingualism (i.e. pronominal systems of English and Spanish); and 3) how Spanish-English bilinguals have been shown to acquire pronominal clitics and clitic-climbing based on their age of exposure to bilingualism.
3 The Language of Bilingualism

3.1 Introduction

The aim of this chapter is to provide some background regarding the unique linguistic behaviours of bilingual speakers in development and at mature stages of knowledge, and to consider how we can account for these behaviours systematically using the language model outlined in Chapter 2. Further, I will provide systematic definitions of terminology used in the field of bilingualism and tease apart the cognitive and experiential aspects of the systematic ways that bilinguals use the language system to satisfy their communicative needs. This background will help to define more precisely how we expect bilinguals to deal with grammatical optionality, and more specifically Spanish clitic-climbing. The chapter will, thus, consider the following topics:

1) The notions of autonomy and interdependence in bilingualism and the language-internal and language-external factors shown to affect bilingual language behaviours;
2) A typology of bilingual language effects and delimitation of their use;
3) A review of cognitive proposals regarding bilingual language behaviour and a systematic approach to the bilingual effects described in the current study.

3.2 Language and Bilingualism

3.2.1 Language Autonomy and Interdependence

Evidence of the unique patterns of language knowledge and use exhibited by bilinguals for developmental and mature grammars provides a good source of data to theorize how the different components of language-related cognition develop and interact. Specifically, in modelling a complete language system that is simple and universal, one must be able to account for evidence of both language autonomy and interdependence in the developmental and mature grammars of child and adult bilinguals. Further, one must also be able to account for how a range of language-internal (e.g. language typologies) and language-external factors (e.g. language dominance) affect the types and degree of autonomy and interdependence exhibited by different populations of bilingual speakers. Before considering these factors, however, I will define more precisely what is meant by language autonomy and interdependence.
In the current study, language autonomy means that a bilingual speaker exhibits patterns of knowledge and use—quantitatively (rate) and qualitatively (form)—for a linguistic expression that converges to that which a monolingual speaker would exhibit in development or as a mature speaker (Paradis and Genesee, 1996; White, 2003). In this case, the bilingual speaker has been exposed to adequate primary linguistic data in each language (by any particular stage of development) to acquire monolingual-like knowledge of the grammatical domain, is able to process this knowledge in a language-specific way, and is able to effectively inhibit the cross-linguistic activation of language knowledge. Consequently, the bilingual is able to strictly use the language resources of the target language, and block out the linguistic resources of the non-target language, in the communicative practice demanded of them.

Conversely, in the case where bilingual speakers exhibit non-monolingual-like knowledge or use of a particular linguistic expression (in development or at mature linguistic stages), there must be a way that dual-language exposure affects the speaker’s analysis and/or construction of that expression. I consider four possible types of language interdependence here: 1) code-switching; 2) transfer; 3) cross-language influence; and 4) general bilingual effect. While the first three can be considered bilingual effects, here I will make the distinction between a bilingual interaction effect and a general bilingual effect. In the case of a general bilingual effect, a bilingual speaker does not use cross-language resources (directly or indirectly) to construct a linguistic expression, but rather, uses universal or unmarked linguistic features at any stage because of altered patterns of exposure to the target language (as a result of bilingualism). If the unmarked/universal option is characteristic of the adult target grammar, but distributed differently than attested by bilinguals, then an apparent quantitative effect will emerge. Conversely if the universal/unmarked structure is not characteristic of the adult grammar, an apparent qualitative effect may occur.

On the other hand, the three types of bilingual interaction effects considered here (code-switching, cross-language influence and transfer) emerge when the linguistic resources of the non-target language affect (directly or indirectly) the formulation of a linguistic expression in the target language. Code-switching and transfer occur when bilinguals directly use the language resources of both languages directly to construct a linguistic expression. In the case of code-switching, speakers draw upon both grammatical and phonological features of both languages in
a highly constrained way. For transfer, bilinguals draw upon both grammatical and/or phonological resources of the non-target language; however, in expressing the utterance, the speaker will attempt to express him or herself exclusively in the target language (and not switch codes mid-utterance), often resulting in a linguistic expression that diverges qualitatively from target norms. For cross-language influence, however, bilinguals only draw upon target-language resources, but the relative availability/activation of these resources is altered by representations of non-target-language; as such these types of interaction effects are indirect and mainly quantitative. One representation of the target-language is over/underused as compared to monolinguals because the non-target language provides strong evidence of a particular word order/grammatical representation, enhancing the activation of the representation in the overall language system.

While a general bilingual effect typically results from having different patterns of exposure to the target language that lead to divergence from monolingual norms at any particular developmental moment (in development and at mature stages), bilingual interaction effects can result from altered patterns of exposure to the target language, or from situational factors where a bilingual recognizes a need/opportunity to communicate in a more bilingual-like way. In the following section, I will further consider these two conditions—functional and contextual factors—that motivate non-monolingual-like linguistic behaviour in bilingualism.

3.2.2 Functional and Contextual Bilingual Effects

In general, there appear to be two main reasons why a bilingual speaker exhibits language effects not exhibited by monolingual speakers:

1) Due to a functional need the speaker has to use universal or cross-language resources for the analysis or construction of a linguistic expression in one language;

2) Due a situational need/opportunity the bilingual identifies in which he or she can communicate like a bilingual; he or she either uses a more efficient processing strategy or communicates in a way that is relevant to a particular speech group (related to group identity, sociocultural factors, etc.).

In the first case, bilinguals may often be unable to maintain the use of strict language-specific resources because they either lack the linguistic knowledge needed to analyze or
construct a particular linguistic expression, or they have some sort of difficulty in processing that knowledge in the most language-specific way. When a bilingual lacks a language-specific linguistic representation (property or features instantiated for property) in one of his or her languages (common to L2 speakers and unbalanced bilinguals), he or she will often directly use the resources of the other language (L1 or stronger/more reliable language) to fill in the morphosyntactic gap. However, a bilingual may exhibit language-specific knowledge of a morphosyntactic domain, but fail to express this knowledge reliably in production or on other performance-related tasks. As Kroll et al. (2006) point out, bilingual language effects may be the result of a speaker not using strict language-specific resources at a variety of areas (i.e. loci) of the language system. Consequently, non-monolingual-like behaviour is not necessarily the result of one unique issue that bilinguals have in keeping language resources independent of each other, but rather may be the result of a (temporary or stable) need the bilingual has to bolster semantic, morphosyntactic or morphophonological resources (for any particular domain) in the target language. In terms of functional effects, therefore, bilinguals may have either a representational deficit or a processing demand that leads to their recruitment of alternative resources (other than the resources of the target language) to fulfill their communicative needs. The resultant linguistic expressions (in production), therefore, can range from strong qualitative effects which affect the overall intelligibility of the utterance, as in (39), or may be more subtle quantitative effects, where the gross intelligibility of the utterance in unaffected, but is obviously non-monolingual-like, as in (40).

(39) Where’s the Santa Claus give me the gun? (Yip and Matthews, 2000, p. 204)
(40) I call Bill this morning and nobody answer. (Lardiere, 2006b, p. 405)

Even though the speaker may be able to identify that he or she is not using, or has not used, the correct structure or form to express his or her communicative intent, the process by which non-target-language resources are used in the construction (or analysis) of a linguistic expression for functional purposes occur automatically, even though the goal of the speaker may be to communicate in the most language-specific way possible.

In the second case, bilinguals may often identify contexts in which the use of strict language-specific resources is not needed or preferred for a particular communicative act. Here, Grosjean’s (1998; 2001; Soares & Grosjean, 1984) notion of language mode is relevant for the consideration of contextual bilingual effects. In situations in which a (typically high-functioning)
bilingual identifies that he or she is communicating in a bilingual context, he or she may (mostly unconsciously, but sometimes on purpose) shift into a bilingual mode, and consequently exhibit non-monolingual-like language behaviour that ranges from subtle effects in single-language utterances (quantitatively or qualitatively different from a monolingual) to more obvious alternations between linguistic codes (code-switching). Here, I will not consider code-switching at length\(^\text{24}\); however, for single-language utterances, the non-monolingual-like behaviour may also range from more subtle quantitative effects (indirect effect from non-target language) to more obvious qualitative effects, where it appears as though the bilingual is lexicalizing syntactic information from the non-target language. Being in a bilingual mode, therefore, may result from either a deliberate act on the part of a bilingual to communicate like a bilingual, or may be a more efficient way a bilingual can communicate when he or she realizes that his or her interlocutor(s) shares the overall language knowledge that he or she has. As such, in a bilingual mode the speaker may relax whatever mechanism is responsible for maintaining a strict separation between language-specific resources in the construction of a linguistic expression. Whatever the contextual motivation, being in a bilingual mode may have consequences for:

1) the selection of linguistic information in the construction of a linguistic expression (selection of morphosyntactic features, etc.);
2) the relative use (more?) of similar or shared cross-language forms/representations;
3) long-term grammatical effects where bilinguals converge upon a new representation for a particular grammatical domain (i.e. convergence of Sánchez, 2004; Toribio, 2004).

I will review the systematic details of language selectivity in section 3.3.1 and consider bilingual convergence in more length in Chapter 4. Further, in section 3.3.2, I will review the empirical work that shows that the different types of bilingual language effects, including cross-language influence and transfer, may occur for both functional and contextual purposes in early and late bilinguals.

---

\(^{24}\) I do not go into great detail in systematically describing how code-switching works in the language model outlined in Chapter 2 because it is outside of the scope of inquiry for the current study. However, as others have pointed out in their investigation of code-switching from a linguistic perspective (Poplack, 1980; Myers-Scotton, 1993), this type of bilingual interaction effect is highly constrained from a linguistic point of view. It is a type of interaction effect that occurs for proficient bilingual speakers who are able to restrict their alternations according to universal and language-specific rules.
In sum, there are two basic reasons why bilingual effects may occur in the language practices of bilingual speakers: 1) for functionally-motivated purposes; and/or 2) for contextually-motivated purposes. In the first case, any group of bilingual speakers may have incomplete language-specific knowledge (representations) for a grammatical domain or difficulty in processing that knowledge in the formulation of a particular linguistic expression. Consequently, they may use the universal/non-target language resources to fulfill grammatical or lexical needs regardless of the context of communication (bilingual or monolingual mode). In the second case, more fluent bilinguals may have control over the degree to which they use cross-language resources to construct a linguistic expression. This expression may be a well-constrained alternation between the two linguistic codes (code-switching) or may be a well-constrained use of lexical or grammatical resources from the non-target language that are interpretable in the target language for a bilingual context of communication. Nevertheless, regardless of whether the bilingual effect occurs for functional or contextual purposes, the types and degree of bilingual effects occur on the basis of many language-internal and language-external factors, such as (but not limited to) the typology of language-pairings, the relative level of proficiency in the target language and control over the speech act. These language-internal and language-external factors that affect bilingual language behaviours are examples of a variety of variables that may contribute to asymmetries in the ways that bilingual speakers develop, represent and use language knowledge as compared to monolinguals, and also to asymmetries in the way that bilingual speakers of different ages build, represent and use language knowledge. In the next two sections, I will consider aspects of cognition (language-internal) and experience (language-external) that have been shown to affect the types and degree of bilingual language behaviours exhibited in child and adult bilingualism.

3.2.3 Language-Internal Factors of Bilingual Language Behaviour

In general, there are two main cognitive asymmetries that make a context of dual-language exposure different from that of single-language exposure:

1) bilinguals must represent and organize language resources from two languages and find ways to enhance or inhibit one language to express themselves in the most (efficient) language-specific way in the other; and
2) bilinguals must know what areas of grammar are typologically compatible cross-linguistically so that the appropriate non-target-language resources can be used (or eventually avoided) for an expression in the target language when bilingual interaction effects do occur.

In the first case, all bilinguals, regardless of age, must be able to analyze or build linguistic expressions in the most language-specific way. With the language model used here, the role of a basic inhibitory mechanism in bilingualism is responsible for reducing or eliminating the activation potential of lexical or grammatical features (mostly incongruent cross-linguistically) not belonging to the target language (horizontal inhibition). This mechanism does not alter the activation levels of two competing variants (with grammatically congruent features) on the basis of declarative knowledge associated with the particular communicative act (vertical inhibition). Nevertheless, as outlined in the previous section, some bilinguals (i.e. highly-fluent) may use declarative knowledge related to relevant situational variables in order to make decisions on how to formulate a particular linguistic expression (i.e. knowing what is interpretable or not for another bilingual interlocutor). These bilinguals, therefore, appear to have both a basic (horizontal) inhibitory mechanism to reduce or eliminate cross-language activation, as well as a more controlled (vertical) mechanism by which they can use declarative information to make sociolinguistically-informed choices (both intra- and cross-linguistically) during monolingual and bilingual modes of communication.

In section 3.3.1.2, I present empirical work that supports the notion that there is some cognitive function responsible for the suppression of cross-language resources in bilingualism, at least for (most) speakers who are not highly fluent and (equally) functional in both languages. For bilinguals who are not as fluent and/or functional in one language, who are at early stages of development (child 2L1 or adult L2), or who need a more economical way to formulate a linguistic expression at any given moment, this cross-language inhibitory mechanism may not be maximally effective, or may vary according to the communicative needs of the particular bilingual. Why any particular speaker will exhibit a general bilingual effect or a bilingual interaction effect for a given grammatical domain (and not others) appears to be affected by a number of factors, including the way that the languages are typologically similar for that particular grammatical domain.
This second cognitive-related factor associated to contexts of dual-language exposure—the ability of the bilingual to automatically know what language resources can be used cross-linguistically—has been proposed in a variety of accounts to be one of the basic factors that influences whether a bilingual interaction effect (i.e. use of cross-language resources) will occur for any grammatical domain (Tracy, 1995; Rothman, 2009), or alternately influences the way in which bilinguals process particular grammatical domains compared to monolinguals (Zawiszewski et al., 2011). This particular variable is commonly known as the typological similarity between languages, and has been referred to in certain accounts in terms of language overlap or grammatical/structural overlap (Muller and Hulk, 2001). While most accounts don’t systematically specify what is meant by typological similarities, here I will describe two ways to consider this notion from a grammatical point of view: 1) according to word order (surface structure); and 2) according to morphosyntactic features (underlying syntax). Empirical evidence from studies on early and late bilingualism generally show that cross-language similarities in word order (surface structure) may lead to two different types of bilingual effects:

1) a quantitative effect where the speaker overuses a certain word order/item in the target language with no interpretational consequences;

2) a quantitative or qualitative effect where the speaker overuses a certain word order/item in the target language for a given interpretation or uses a word order/item in a way not found in expressions constructed by monolingual speakers.

An example of the first type of effect comes from evidence of word order variability in the production of intransitive sentences. It is well-known theoretically that there are two types of intransitive verbs: unergatives (dance, sneeze, etc.) and unaccusatives (so-called presentational verbs—arrive, enter, etc.) (Perlmutter, 1978; Burzio, 1986; Levin and Rappaport-Hovav, 1995).

---

25 Zawiszewski et al. (2011) recently examined the processing of Basque by adult Spanish-Basque bilinguals (age of first exposure to Basque: 3;0) and showed differential processing (ERP measures) between these early bilingual speakers and native speakers of Basque and Spanish depending on the congruence of underlying syntax between Spanish and Basque. When the grammars were incongruent (i.e. head parameter and ergative/accusative syntax), bilingual speakers processed the grammar in a different way than native speakers, while there were no processing differences when the grammars were congruent (i.e. verb agreement). The authors suggested that language distance is a factor that can affect the way that these (early) bilinguals process language as compared to the native speakers tested in their study. Whether these results are relevant to cognitive factors associated with age of first exposure or with bilingualism in general has yet to be determined, as it was not indicated in the study if the native speakers were monolingual or not at the time of testing.
In the case of Italian and Spanish, there are a variety of ways in which this difference is linguistically encoded, including through word order variation (SV and VS) in unmarked (neutral subject) declarative sentences and, for Italian, through a variation in the use of auxiliaries in the formation of the periphrastic perfective past tense. English, on the other hand, does not encode this difference through word order variation, nor through the use of perfective auxiliaries.

Examples of these differences are illustrated in (41)-(43).

(41)  i. Mi nieto gritó. [unergatives: SV, Hertel, 2000, p. 2]
     my grandson yelled
     “My grandson yelled.”

to. Llegó un hombre. [unaccusative: VS, Hertel 2000, p. 19]
     arrived a man
     “A man arrived.”

(42)  i. I professori hanno parlato a lungo. [unergatives: avere, Sorace, 1995, p. 155]
     the teachers aux-have talk-pp at length
     “The teacher (have) talked for awhile” (my translation)

     many tourists aux-are leave-pp
     “Many tourists (have) left.” (my translation)

(43)  i. My grandson yelled/A man arrived. [intransitives: SV]

    ii. I (have) talked a long time/I (have) left. [periphrastic perfective past: have]

Evidence from studies on Italian-English and Spanish-English bilinguals show that even though bilinguals appear to acquire the basic syntactic difference between the two types of intransitive verbs (Sorace, 1995; Hertel, 2000; Lozano, 2006), they may still exhibit patterns of knowledge and use that diverge from the most language-specific patterns exhibited by native/monolingual speakers (especially see Sorace, 1993, 1995). For example, in a production task, Hertel (2000) found that low to high intermediate English speakers of L2 Spanish used the SV word order for both unaccusative and unergative verbs in unmarked declarative contexts (as responses to global questions) almost categorically (>90%), while native and advanced L2 speakers produced significantly higher rates of the VS word order with unaccusative verbs (39% and 55%, respectively) as compared to unergatives (6% and 33%, respectively). In a judgment task, however, all but the low proficiency group exhibited some knowledge of the unergative/unaccusative distinction, exhibiting preferences for the VS word order more for
unaccusatives than for unergatives (convergence to native like judgments increased with proficiency). These results suggest that there is something about their experience with bilingualism (i.e. strong evidence of SV from non-target language, relative patterns of exposure to the target language) that makes certain bilingual groups rely on the SV word order more for the production of intransitive verbs in these particular contexts.

Bilingual speakers, therefore, can use cross-language resources to resolve ambiguities related to word order variations in a target-language even if they appear to have acquired the underlying syntax for the grammatical domain. Nevertheless, in most cases where there is word order variability (surface structure) in a language, there is also some sort of interpretational consequence, captured in basic linguistic principles of blocking effects (see review in Embick, 2008). Nevertheless, this does not mean that all semantic features are expressed in ways that are transparent in surface structure; here we are reminded of the logical problem in generative-based acquisition studies: How is it that speakers come to know the formal knowledge of language (underlying features) through experience with surface structures alone? In the case where a speaker has to solve a learnability problem through experience with input (surface syntax), and that input is variable (word order variations), bilinguals have an additional source of information that monolinguals don’t have: knowledge (developing or mature) of the non-target language. Consequently, if the (invariable) surface structure of the grammatical domain in the non-target language is similar to that of a word order variant in the target language (thus providing strong evidence to resolve the word order problem), the bilingual speaker may use that word order more in the target language with the concomitant interpretational consequences. Two such examples of this type of bilingual effect comes from evidence of the use of overt subject pronouns and from evidence of the use of direct object marking (DOM) by Spanish-English bilinguals.

In most Spanish dialects the use of overt subject pronouns is marked from a semantic-pragmatic point of view: as a way to focus the subject constituent (for contrast, emphasis, switch-reference, etc.), as in (45). In unmarked discourse contexts, the null subject is the preferred variant, as illustrated in (44). English is a non-null subject language, and as such, speakers must use an overt subject pronoun (SV) in both marked and unmarked discourse contexts (and use prosodic cues to focus the marked structure).
i. Q. ¿Qué pasó?
   A. Ø Necesito ayuda. [unmarked, null subject]
      Ø need-1.sing. help
      “I need help”

   ii. Q. ¿María necesita ayuda?
      A. No, yo necesito ayuda. [marked, overt subject]
         no I need-1.sing. help
         “No, I need help.”

(45) i. Q. What’s up?
   A. I need help. [unmarked, overt subject]

   ii. Q. Does Mary need help?
      A. No, I need help. [marked, prosodic cue]

In general, Spanish-English bilinguals are quite successful at acquiring the underlying syntax of null subjects (White, 1986; Phinney, 1987; Liceras, 1989; Al-Kasey & Pérez-Leroux, 1998; Lozano 2002); however, empirical evidence has shown that bilinguals may overgenerate overt subject pronouns in unmarked contexts (Al-Kasey & Pérez-Leroux, 1998; Montrul 2006) and apply reference inappropriately to overt subjects in embedded contexts (Pérez-Leroux & Glass, 1999). The surface structure of English, therefore, may affect the way that bilinguals often encode semantic information to pronominal subjects and/or process their lexical choices differently than monolinguals, leading to a bilingual effect with subtle interpretational consequences.

Similarly, in Spanish there are two ways to linguistically encode a direct object when it is animate: with the direct object marker a if the direct object is [+specific] and no marking when the direct object is [-specific]. Inanimate direct objects do not exhibit any marking for specificity in any context. In English, the grammatical and semantic information encoded to the Spanish ‘a’ is recovered from word order and the context of discourse, and as such, all direct objects in English are expressed like Spanish inanimate and non-specific animate direct objects (examples from Montrul & Bowles, 2009, p. 365):
Montrul & Bowles (2009) examined the knowledge and use of direct object marking by early (heritage) Spanish-English bilinguals and showed that adult bilinguals of all proficiency levels appear to have indeterminate knowledge of this domain as shown by probabilistic judgments and production of DOM (and other relevant structures related to inherent case). Again, the bilingual effect leads to a subtle difference between the bilingual and monolingual grammars, where an apparent simplification of the bilingual system occurs on the basis of altered patterns of input in bilingualism (strong evidence of a certain word order pattern cross-linguistically and relatively insufficient evidence in Spanish of the linguistic encoding of [+specific +animate] objects).

Here, I have given two examples where Spanish-English bilinguals may be able to use the surface structure of the non-target language (English) in an attempt to resolve a learnability problem associated with word order variability in the target language (Spanish). Crucially, however, Spanish speakers can grammatically analyze overt subjects and specific, animate objects without a, even if they’re encoded with English-like (or unmarked) features; the underlying syntax of Spanish overt subjects and direct objects is relatively congruent to English (XPs). Nevertheless, there are many cases in which the underlying syntax is not congruent cross-linguistically, such as in the case of object pronouns in Spanish (and other Romance languages) and English: in Spanish object pronouns are clitics (head constituents, X₀) and in English they are phrasal constituents (XP). In this case, where the underlying syntax of the two languages is not congruent, the bilingual may do one of two things:

1) resort to universal or unmarked features for that grammatical domain to resolve the ambiguity; or
2) force the use of incongruent features from the non-target language, resulting in an expression that is usually badly-formed (qualitatively) from a grammatical point of view (i.e. lexicalizing the syntactic frame of the non-target language).

For example, object clitics may present a learnability challenge to speakers of a Romance language because of the variability that characterizes their target distributions (pre- or post-verbal, i.e. Spanish) and/or relative use in input (i.e. frequent omissions, i.e. French). Bilingual speakers of a Romance language and English, however, have a strong source of information favouring the use of (marked) strong object pronouns in post-verbal position (Romance languages also have strong object pronouns, XPs). However, given that the underlying syntax of unmarked object pronouns is incongruent cross-linguistically, the question is whether or not bilinguals will utilize cross-language knowledge of strong pronouns to resolve the learnability problem of Romance clitics, at least for a while. Evidence from acquisition studies points to age effects here: early bilinguals of a Romance language and English (simultaneous, early child L2), have been shown to resort to universal/unmarked syntactic features (Pérez-Leroux, Pirvulescu & Roberge 2009; Pérez-Leroux, Cuza & Thomas 2011a), while late child and adult bilinguals often exhibit early syntactic misanalyses by using XP syntax with unmarked object pronouns (clitics) until they have figured out the underlying syntax of Romance clitics (Liceras, 1985; Grondin & White, 1996). For example, Pérez-Leroux, Pirvulescu & Roberge (2009) showed that early bilingual French-English children persist in their use of null objects in French, as in (48), past the stage at which their monolingual counterparts begin to use clitics reliably (>90% of obligatory contexts). Nevertheless, English speakers of late child or adult L2 French may initially analyze the pronoun as an XP (like English) and exhibit a bilingual interaction effect (post-verbal pronouns with morphophonological shape of clitics) not exhibited by monolingual speakers at any point of their development, as in (49) (Selinker, Swain & Dumas, 1972; Grondin & White, 1996):

(48) J’ai déjà lu. [target: Je l’ai déjà lu] 
I aux-pst. already read-pp 
“I’ve already read/ I read already.” [target: I’ve already read it. (asked about a book)] 
(example in Pérez-Leroux, Pirvulescu & Roberge, 2009)

(49) Le chien a mangé les. [target: Le chien les a mangé.] 
The dog aux-pst eat-pp pron-3-pl. 
“The dog ate them.” 
(example in Selinker, Swain & Dumas, 1972)
In building grammars with UG as the starting point, early child bilinguals appear to maximally differentiate grammars that are incongruent in underlying syntax; conversely, late child and adult bilinguals build a new grammar with their L1 (and in principle, UG) as the starting point, and as such, may analyze the morphophonological features of the target language with the underlying syntax of the L1 until they have enough evidence to identify the lack of syntactic congruence (at which point they may acquire properties not instantiated in L1 through UG). As the data above suggests, therefore, a variety of language-external factors may influence whether a general bilingual effect or a bilingual interaction effect occurs, such as age of first exposure to bilingualism, and patterns of exposure to the target-language (i.e. quantity, quality and mode of input). I will consider these factors, as they influence bilingual language behaviours, in the following section.

3.2.4 Language-External Factors of Bilingual Language Behaviour

Similar to the notion of typological similarity, the notion of language dominance in bilingual studies is used frequently, but rarely defined from a systematic point of view. Generally, it appears as though language dominance, as an external factor affecting bilingual language behaviour, is interpreted from a relative fluency point of view: a bilingual is more or less dominant in one language because he or she exhibits better fluency (knowledge and use) in one grammar (or grammatical domain) over the other. Again, what this means from a systematic point of view is unclear; I will leave this point for now and return to it shortly.

Dominance, as a linguistic variable, subsumes a number of experiential factors related to language exposure: age of exposure, quantity of exposure, quality of exposure and mode of exposure. In the first case, it may seem obvious as to why acquiring a second language as an older child or adult may lead to L1 dominance effects for any linguistic domain, especially those that present learnability challenges in the L2. Age of exposure to bilingualism appears to affect how language dominance functions as a linguistic variable, especially at the early stages of language development. The process of becoming bilingual for an adult is to start off as a very imbalanced bilingual (strongly dominant in L1), while the process of becoming bilingual for a baby/young child is to start off as a balanced bilingual with UG serving as the starting point for
both grammars. Nevertheless, acquiring two languages as a native speaker (from birth) does not guarantee that this speaker will maintain a language balance and converge to monolingual-like patterns of knowledge or use (overall or for any particular domain) in both languages. Cross-language asymmetries often characterize the exposure to input that early bilinguals have for each language: asymmetrical quantities of input in each language, asymmetrical qualities of input in each language (monolingual speakers, other bilingual speakers) and asymmetries in the type of exposure they have in each language and in their opportunities for communicative interaction (education, literacy, exposure to a variety of registers, speakers, etc.). These asymmetries also characterize input-related differences between early child bilinguals and late child and adult L2 speakers of a language, and consequently, may affect the types of bilingual effects that emerge in development and at mature stages of language knowledge for different groups of speakers.

Language dominance, therefore, can be viewed from two different perspectives: 1) in the asymmetrical acquisition of linguistic knowledge for any grammatical domain; and 2) in the asymmetrical acquisition of sociolinguistically-relevant knowledge that may affect the relative completeness of a bilingual’s language knowledge for any grammatical domain, or for overall fluency. In the first case, the types of bilingual effects that occur are mostly due to the way that bilinguals use overall core linguistic knowledge to analyze or build a linguistic expression. As described above, bilinguals, unlike monolinguals, may use either cross-language resources or universal resources (universal or unmarked features) to resolve a learnability problem if they cannot use the target-language resources to do so. If a bilingual is able to identify that the underlying syntax of one language is incongruent with the other, then he or she may resort to unmarked or universal features; this type of bilingual effect is common to early bilinguals, and may appear as a delay in the bilingual speaker’s acquisition of a particular grammatical domain, although, as Francis (2011) points out, this is not a pathological delay in any sense. This type of effect is also typically called a quantitative effect because the use of universal features persists in the developing grammar of the bilingual longer than for the monolingual; as such, when a bilingual is compared to a monolingual at the same age/stage of development (often measured by (w)MLU or other linguistic measures), it appears as though the bilingual uses less of a particular feature/structure than the monolingual (i.e. clitics in French).

Nevertheless, if the bilingual exhibits a very strong asymmetry in his or her knowledge of the linguistic properties/features of the two languages, then he or she may erroneously use cross-
language resources to resolve a learnability (or communicative) issue. Here we may see quantitative effects (more or less of a certain structure compared to a monolingual) or may see qualitative effects, where a bilingual will produce a structure in the target language that a monolingual would never produce in development or at mature stages of knowledge. Yip & Matthews (2000) show such an effect where a young Chinese-English bilingual, who the authors assess as being dominant in Chinese, transfers his knowledge of the syntax of Chinese relative clauses (right-headed) to English at a stage of development (2;05) when he has not yet acquired (left-headed) English relative clauses:

(50)  Where is the Santa Claus give me the gun?  (Yip & Matthews, 2000, p. 204)
     Where is the gun Santa Claus gave me?

Yip & Matthews proposed that many of the bilingual language behaviours that this child exhibits in English are influenced by being linguistically dominant in Chinese; evidently, for this bilingual child the expression in (50) is completely interpretable, given the overall language knowledge that he has so far. As such, this type of bilingual interaction effect may be seen as a sort of communicative boosting effect, where a bilingual child has resources that a monolingual child does not have to produce certain structures that emerge asynchronously in language development cross-linguistically. Chinese relatives emerge in monolinguals early in the third year, while English relatives emerge in monolinguals relatively later; this bilingual is able to use cross-language resources, therefore, to say something in English in a way that, normally, he or she wouldn’t be able to express (object relatives), and only does so because he is dominant in Chinese and doesn’t identify that the cross-language structure results in an extremely badly-formed linguistic expression (completely uninterpretable for any English speaker).

Whether this type of interaction effect has a bootstrapping function for bilinguals (as proposed by Gawlitzek-Maiwald & Tracy, 1996) is unclear, as no empirical work has provided strong evidence to suggest that bilinguals may converge to target-like knowledge for a particular grammatical domain in one language earlier than would a monolingual speaker because of the availability of cross-language resources. In fact, young bilingual children do not typically produce qualitatively divergent structures, such as (50) above. Most research has shown that early child (2L1, very early L2) bilinguals know where underlying syntax is incongruent cross-linguistically and avoid using cross-language resources in these cases (Paradis & Genesee 1996;

Given that the Yip & Matthews (2000) type boosting effect seen during early child bilingual development has been seen in research with other child bilinguals, such as in Dopke (1998; 2000) and Gawlizek-Maiwald & Tracy (1996) for German-English bilinguals, the question is, what is it about the experience these bilinguals have with their two languages that leads to these types of bilingual language behaviours that are divergent from a monolingual pattern of development? Most accounts suggest that this type of apparent linguistic imbalance in early child bilinguals emerges as a result of an asymmetrical quantity of exposure to each language: the child is exposed to one language for more of their waking hours than the other language. In cases of unbalanced bilinguals, a child may appear to be dominant in one language at any moment of development because he or she has been exposed to that language for more waking hours than the other language, and as such, has met the input threshold for a particular grammatical domain in his or her stronger language before he does so in the weaker one. This may also be true of late bilinguals, who require enough quantity of input to realize the grammatical incongruencies between their L1 and the target language, and acquire the language-specific features of their L2. As such, late bilinguals range from being strongly L1-dominant bilinguals in the early stages of L2 development and advance to becoming more balanced bilinguals with more exposure to the L2, and as such, are able to acquire the language-specific features of the target language and maximally differentiate between their two grammars.

The trouble with using quantity of input/exposure as a unique deterministic linguistic variable for relative language dominance, however, is that it is difficult to determine how much input/exposure is enough (what the threshold is) for any particular grammatical domain in any particular language given different contexts of acquisition. Consequently, early bilinguals, in principle, can appear to be dominant in one area of grammar, but have had relatively equal quantities of exposure cross-linguistically, or can have had very asymmetrical experiences in terms of quantity of exposure cross-linguistically, but not exhibit language dominance effects for an area of grammar because one of the languages requires a relatively lower threshold of input than the other to acquire a particular domain. Further, it seems improbable that an adult of a late-acquired L2 would ever reach a level at which he had received equal quantities of exposure in
both languages through his lifespan; nevertheless, some adults are able to acquire L2 proficiency levels that are so near-native that other native speakers have a hard time distinguishing their language background (Birdsong, 1992). The relationship between quantity of input and language dominance is, therefore, not as deterministic as is often assumed. As such, quantity of input should be used as one of a number of experiential factors that can lead to bilingual language behaviour as a result of apparent language dominance.

Besides the quantity of input, the quality of input and mode of input can affect the degree to which a bilingual appears to be linguistically dominant in one of his or her languages at any stage of development (in childhood or adulthood). For example, child and adult L2 speakers are often exposed to a range of language speakers who are monolingual, bilingual, native and non-native speakers of the target language. If a speaker is exposed to input that offers strong evidence of a particular bilingual-type speech pattern (i.e. overt subjects or SV word order in unmarked declarative contexts in Spanish), then there is no reason that this speaker should acquire a grammar where strict language-specific (monolingual-like) patterns of use are instantiated for that particular grammatical domain. As such, the speaker will always appear to be dominant in their L1 because they have acquired an L2 grammar that already is divergent from that of monolingual speakers (see Hulk & Cornips, 2006). Recently, researchers examining heritage learners of Spanish in the US have also pointed out that the input provided to US-born Spanish heritage speakers may already be divergent (non-monolingual-like) from a linguistic point of view (Rothman, 2009; Domínguez, 2009). Consequently, heritage speakers (and non-native speakers) may be exposed to primary input characterized by bilingual language behaviour, and as such, may instantiate their Spanish features in a way that is different from monolinguals. In turn, these bilinguals will appear to be English-dominant because of the degree to which it appears that English (or unmarked) features permeate their speech practices in Spanish.

In a similar way, if an L2 speaker has been exposed to formal classroom learning, the bilingual language behaviour may be reinforced through explicit learning (i.e. that it’s OK if they use overt subject pronouns, SV word order). The mode of input is also a factor that can affect the

---

26 This particular type of bilingualism has often been compared to a creolization process. See the papers in Lefebvre, White & Jourdan (2006) for a more detailed presentation of such work.
apparent language dominance of a bilingual speaker: they may receive formal instruction that reinforces bilingual language behaviour or, conversely, they may be exposed to a more restricted type of input that reinforces certain bilingual language behaviours typical of that mode of input. Heritage speakers, for example, may exhibit bilingual language behaviours that are typical of the vernacular of their speech community. They are often limited in their exposure (i.e. naturalistic input from parents), and as such, may not develop patterns of fluency in their home (heritage) language that they develop in the majority language (i.e. expanded vocabulary, more complex structures, literacy, etc.), thus leading to apparent language dominance in the majority language. For example, despite receiving a robust quantity of exposure in the heritage language by age 5, the vocabulary of the child who learns a minority language at home may always be limited to referents in the home and family environment. If the child has received relatively equal amounts of exposure in the majority language (through daycare involvement, school, neighbours, etc.) the child may have a more expanded vocabulary because of the diverse contexts in which he or she has experience with the majority language. In this case, quantity of exposure to input does not necessarily determine how dominant the child is from a lexical (and consequently, grammatical) point of view, but rather the context of exposure may lead to a more expanded vocabulary (and possibly more complex structures) in one language. While no known empirical study has examined this issue directly, evidence from studies such as Golberg, Paradis & Crago (2008) suggest that the context of input may have a significant impact on the rate of lexical development of bilingual children, especially in majority-minority language contexts. These authors showed that among child learners of L2 English (mean age of exposure 5;04), where English was the majority language of their community, those who had exposure at home (after the start of school) in addition to full day school did not exhibit any advantages over those who only had exposure in school in terms of their rate of lexical development. Here, being exposed to English in a school environment provides sufficient input (quantitatively and qualitatively) to facilitate the rapid learning of L2 vocabulary. Being formally educated in a language, therefore, may affect how apparently dominant a bilingual is (or becomes), regardless of the quantity of time they spend being exposed to any particular language.27

27 Schweiter & Sunderman (2008) show that lexical robustness may be a variable that contributes to relative language proficiency in bilingual populations; if formal education enhances the lexical robustness of one language
Further to strictly linguistic issues, language dominance may also be characterized by a particular speech group’s ability to communicate in a number of different sociolinguistically-relevant ways, depending on factors such as the formality of the speech act, the topic of conversation, the linguistic background of their interlocutors, etc. As such, even if a bilingual is assessed to be more or less balanced in their particular linguistic abilities (i.e. knowledge of grammatical properties/features), they may be unbalanced in their abilities to change their speech practices according to a number of sociolinguistically-relevant variables. As such, a heritage bilingual, for example, may not know how to change his or her language practices to accommodate the type of speech required for written production, thus writing the way he or she would speak. Further, he or she may sound odd to monolingual speakers in his or her home country because he or she has not had a range of input in the target language in order to acquire different language registers of the larger speech community. Similarly, a late L2 learner may only have been exposed to a more formal and corrected language register, and as such, may lack the knowledge of vernacular in order to participate appropriately given the particular context of communication. In this case as well, the speaker may have complete knowledge of a particular grammatical domain, but sound odd in a given context because they have a more restricted set of socio-linguistically relevant grammatical options to choose from (i.e. formal, corrected variants in an informal context).

Language dominance, therefore, as a language-external variable can be viewed in two ways: 1) according to specific cross-language linguistic asymmetries (knowledge of a particular grammatical domain); 2) according to cross-language asymmetries in sociolinguistic knowledge.

over another, then it is not surprising that bilinguals develop a certain degree of linguistic dominance in the language in which they are formally educated.

28 Here I am reminded of two pieces of anecdotal evidence offered by heritage speakers of Farsi and Spanish who returned to visit their home country: in both cases the monolingual speakers of their generation thought their vernacular sounded old or out-of-date, belonging to the former generation of speakers (they were told they sounded like old people). These stories remind us that the input that heritage speakers often receive is a vernacular common to a previous generation (their parents); many do not have extensive social interactions with other speakers of their generation (who are monolingual speakers), and as such, may acquire a type of vernacular that immigrated with his or her parents (and continue to be used in the language interactions by heritage speakers).
In both cases, the bilingual speaker may appear to sound odd (non-monolingual-like) in any given context because he or she

i) formulates a linguistic expression in a quantitatively or qualitatively different way than a monolingual; or

ii) formulates a target-like linguistic expression for a particular grammatical domain that is inappropriate given the situational variables underlying the speech act.

In both cases, procedural knowledge will act to determine how the bilingual formulates the linguistic expression; however, not having declarative knowledge pertaining to the different registers of a language may have an effect where the bilingual appears to be less fluent in any particular context of use. When referring to a bilingual who is balanced or who has so-called complete knowledge, therefore, it may be important to indicate what is meant by this: either they have acquired language-specific linguistic features for all the core grammatical domains needed to participate in linguistic communication in each language, or they have acquired equal knowledge of both languages from both linguistic and sociolinguistic points of view. In the latter case, the bilingual should be expected to have linguistic knowledge in both languages related to their performance in any context of communication: formal and corrected linguistic variants (in speech and writing) and informal and casual variants (in speech and writing), and any linguistic features relevant to the registers in between.

### 3.2.5 Variability and Interpretability in Bilingualism

A final consideration that I will make before discussing how we can systematically account for a range of bilingual language behaviour in child and adults is to review the proposals that have attempted to account for i) why adult bilinguals exhibit non-monolingual-like behaviour for certain grammatical domains and not others; and ii) how we can account for the fact that, as adults, early bilinguals (heritage speakers) may exhibit patterns of (variable) knowledge and use in a language that is similar to that of late bilinguals (L2 speakers) for particular domains/language-pairings.

A general view of why early bilinguals may exhibit grammatical variability or incompleteness for certain domains in adulthood adapts a basic premise from the Regression Hypothesis of Jakobson (1941, cf. Montrul, 2008): the earlier-acquired domains are more
immune to language loss while the later-acquired domains are likely to be more vulnerable to loss. Montrul (2008) discusses this premise and shows that while the acquisition of some grammatical domains tends to support this notion, not all early-acquired domains of grammar are immune to bilingual language effects in adulthood. In studies examining the acquisition of aspect (use of preterite and imperfect past tenses in Spanish) and mood (indicative vs. subjunctive), Montrul (2002; 2006) showed that adult heritage speakers of Spanish (Spa-Eng bilinguals) exhibited more robust knowledge of the earlier-acquired feature, aspect, than the later acquired feature, mood. She suggested that since heritage speakers experience a shift in dominance patterns to English around the age of 4-5, they do not acquire a strong representation for Spanish mood, a grammatical domain that continues to develop in monolinguals until past the age of 5. Conversely, in a more recent study examining the acquisition of differential object marking (DOM)—the use of the Spanish personal a—by adult heritage speakers of Spanish, Montrul and Bowles (2009) demonstrated that, regardless of being an early-acquired feature of Spanish (<4;0), heritage speakers arrive at adulthood exhibiting variability in their knowledge and use of this particular domain. These results support Montrul’s (2008) Weaker Language as L1 Hypothesis, which aims to account for the persistent variability in early-acquired bilingual grammars according to one of the following two reasons: 1) the stabilization of developmental/cross-language representations for late-acquired/vulnerable features; or 2) the attrition of early-acquired domains that are most vulnerable to bilingual effects. She presented this hypothesis to counter the Weaker Language as L2 Hypothesis (Schlyter, 1992), which suggested that the weaker language acquired by early (minority/heritage) bilinguals develops as would an L2 in contexts of late acquisition (i.e. L1 mediating acquisition of the L2, etc.).

While it may be more clear why an early bilingual may exhibit variability in adulthood for late-acquired grammatical domains in their heritage language (i.e. due to shifting exposure patterns and stabilization effects), it is still unclear as to why an early bilingual would lose a part of their grammar through childhood and adolescence, and why many of these variable grammatical domains resemble those domains that late bilinguals (adult L2) tend to exhibit variability for at more advanced levels of L2 proficiency (i.e. fail to acquire completely). To account for the apparent selective variability that both early and late adult bilinguals exhibit for certain grammatical domains, Sorace (2005) proposed the Interface Vulnerability Hypothesis to account for a wide variety of data from child and adult language studies that have suggested that
the areas that bilinguals have most trouble with are those that require some mapping or coordination between syntax and semantics/pragmatics (Muller & Hulk, 2001; Belletti, Bennati & Sorace, 2007; White, 2008, among many others). In general, the interface proposal has been quite successful as a descriptive account of persistent variability in bilingual grammars: many studies have found that early and late bilinguals exhibit persistent variability (or residual optionality, Sorace, 1993) in advanced grammars for domains of grammar where the linguistic encoding of semantic features does not happen on a one-to-one basis. That is to say, speakers often have difficulty in figuring out the distribution of certain morphosyntactic items (free or bound morphology) in a language because their distribution is guided by (often subtle) semantic or pragmatic information (overt subjects in null-subject languages, gender agreement, DOM, etc.). As such, proponents of an interface proposal seem to be suggesting that for these bilingual speakers, there is some problem with the functioning of the interface between the formulation of a conceptual plan (the semantics) and the linguistic encoding of that plan (the syntax).

While this explanation has good descriptive power, I am unsure how this explanation is different from the general notion of learnability. If the interface proposal were to have deterministic power to explain the persistent variability in contexts of heritage or adult L2 acquisition, then we need a more clear account of what it is about the interface that is problematic for these speakers: Has it eroded and is no longer able to efficiently coordinate syntactic and semantic/pragmatic information by the time the bilingual arrives at adulthood? Is the flow of information across the interface interrupted because of some synaptic misfire or other neurological or cognitive alteration that occurs in contexts of bilingualism? Or is the flow of information too overwhelming in bilingualism that the interface doesn’t function properly? If this is the case, what information gets through and what doesn’t? How does the system decide?

If the interface is really the problem, then there should be clear empirical evidence not only supporting variability for domains of grammar at the syntax-semantics interface, but also clear and categorical empirical evidence of no variability for domains of grammar where semantic information does not have to be encoded linguistically, or rather where syntax has no consequence for interpretation. Of course, the notion of core syntax relates to properties of grammar, such as head directionality, that tend to be early-acquired features in both child and adult acquisition contexts (and necessary to be speaking language X). While it is difficult to think
of many morphosyntactic items (evident in input) that do not represent the encoding of some semantic information, some recent studies have started to show that, contrary to the predictions of the interface proposal, some advanced bilinguals do exhibit persistent variability for core areas of grammar (Argyri & Sorace, 2007; Iverson, 2009).

Until its proponents define a more clear deterministic model that explains why (and how) the syntax-semantics/pragmatic interface is somehow weakened in contexts of bilingualism, therefore, I will adopt a learnability perspective to explain bilingual language behaviour in the current study, and a dual interpretability proposal to explain why some of that variability persists in adulthood/advanced grammars. The areas of grammar that are going to be most susceptible to variability (in development and at mature stages), therefore, are those where there is something about the input (surface structure) that underestimates what the speaker has to learn about that particular domain. For example, to learn subject syntax in Spanish, English speakers must learn that the alternating presence and absence of an overt subject pronoun (in input) is happening for semantic/pragmatic reasons, and that there are different rules for this variability depending on a variety of language-internal and language-external variables (main vs. embedded context, dialect, etc.). In this case, the speaker has to learn that null expressions (often formalized as empty categories, *pro*, etc.) have some semantic value in Spanish.

Beyond the basic problem of learnability, speakers must also be able to attest to the fact that their own non-monolingual-like knowledge/use of that particular domain (quantitative or qualitative difference) is not the most target-language-specific way (monolingual-like way) to interpret/express that utterance. Here, there are two main ways that persistent variability may occur for adult bilinguals of advanced proficiency. First, the bilingual has been exposed to input (mostly from other bilinguals, but dialectally as well) that supports the hypothesis they’re making about how that particular morphosyntactic element is distributed in the target grammar. As such, a variety of quantitative (and qualitative) effects may persist in advanced (and native) bilingual grammars because they are also typical of the speech variety the speaker has most been in contact with (Paradis & Navarro, 2003; Hulk & Cornips, 2006; Rothman 2007).

Second, where more obvious quantitative (tense omissions) or qualitative effects (resumptive pronouns in English) persist in advanced adult bilingual grammars, a certain
stabilization of that variability may occur. Here, I will suggest that this stabilization may occur because the utterances are interpretable in two ways: 1) the structure is interpretable from a computational perspective (i.e. constrained by UG and language-specific features); and 2) the structure can be interpreted in a target-like way (semantically) for the intended expression in the target language. Why any one speaker will advance away from this variability more than another given similar contexts of exposure (quantity, quality, mode, etc.) may occur for more functional-related reasons. For example, some speakers may have better abilities to exploit the processes of self-monitoring/meta-linguistic processes than others, and in turn, may be able to keep restricting their hypotheses for that particular form/structure until it converges to the most language-specific grammar possible (that they have attested to in input). Nevertheless, not all speakers appear to have the same ability to self-monitor their grammar, and as such, may have variable abilities to perceive their bilingual-like behaviour, and consequently exhibit persistent bilingual effects throughout their lives (see Segalowitz, 2005). Further, for some bilinguals, there is a more conscious need or desire to converge to monolingual-like targets due to individual differences regarding linguistic accommodation, group identity, etc. Basically, some speakers care more than others that their linguistic expressions are correct (i.e. that they sound the most native as possible or not); crucially, these speakers must be aware of the bilingual-like effects in their grammar in order to keep restricting their language hypotheses until they arrive at the most native-like representation (monolingual-like). The proposal of dual interpretability put forth here, therefore, is both formal and functional: the persistent non-monolingual-like behaviour must be interpretable from a formal perspective (constrained by UG, language-specific features), but its persistence may occur due to the fact that these bilingual forms/structures are fully functional from a communicative point of view. The bilingual may not be aware or may not care that the forms/structures do not converge to the most target-language-specific grammar possible for that domain; they may not be aware (or not care) because that particular form/structure allows them to express exactly what they intended to say (in a formal and conceptual way semantically), and as such, there is no need for reanalysis and a further restriction of hypotheses.

Finally, where bilingual-like behaviour persists over time and generations, we may begin, and have begun to see, shifts in the underlying grammar represented by certain bilingual populations. With evidence from Spanish-Quechua bilingualism, Sánchez (2003; 2004) proposed the Functional Convergence Hypothesis to account for contexts in which early bilingual speakers
appear to have instantiated a converged bilingual representation for particular domains, one that exploits linguistic features from both of the languages to which the speaker has been exposed (Silva-Corvalán, 1994; 2003; Toribio, 2004; Montrul, 2004a; Bullock & Toribio, 2006). Here, different factors such as quality of input, context and relative quantity of exposure to both languages may assuage or reinforce such convergence, and it would not be surprising that if most of these domains where young bilinguals instantiate unique bilingual representations are those which present some particular learnability challenge in acquisition. In this sense, bilinguals appear to use their complete language knowledge in the most strategic and ingenious ways, and consequently they interpret and use their language as a monolingual never would (could). In the next section, I will examine these bilingual-like behaviours more closely and use the language model outlined in Chapter 2 to offer a systematic approach to child and adult bilingual language effects.

### 3.3 A Systematic Approach to Bilingual Language Behaviour

In this section, I will discuss how the types of bilingual language effects, and the factors affecting their use, can be systematically accounted for in the language model outlined in Chapter 2. Specifically I focus on describing the nature of the bilingual lexicon and what is meant by a cross-language inhibitory mechanism in a system that employs an activation spreading approach to the organization of linguistic information. Further I will review how the model in Chapter 2 works with a bilingual lexicon to account for both contextual and functional bilingual effects produced by child and adult bilinguals. Finally I will describe three types of single-utterance bilingual language behaviours—general bilingual effect, cross-language influence and transfer—and the systematic ways that these bilingual effects come to be produced by child and adult bilinguals. I assume that a Minimalist-type grammar is the best way to account for these types of bilingual effects because:

i) an invariable grammar can compute a variety of bilingual language forms, as long as they respect universal constraints on linguistic computation; and

ii) these bilingual language forms, even if they diverge from the most language-specific grammar, can be interpreted by speakers of the target language because they (the interlocutors) have the needed language knowledge/resources to do so.
Nevertheless, in many cases bilingual utterances are not interpretable (misunderstood or completely ill-formed grammatically); here, I will discuss what cognitive factors may contribute to the building of quantitatively or qualitatively divergent structures by different groups of bilinguals (i.e. child and adult, low to high proficiency, etc.).

3.3.1 A Bilingual Language System

3.3.1.1 The Bilingual Lexicon

In describing the bilingual lexicon, I will aim to systematically account for the following:

1) The linguistic information encoded in a lexical entry and how lexical entries are the locus of cross-language variability; and

2) The organization of the bilingual lexicon and how language information from one language can be used (or suppressed) for the formulation of a linguistic expression in another.

In the first case, I follow Levelt (1989) (among others) and assume that two basic pieces of information are encoded to a lexical entry: morphophonological information and lexical-syntactic information (the lemma); semantic information is relevant for building both parts of the lexical entry and must be available on both parts of the entry. Further, I will follow generative-based principles (Chomsky, 1965; Pinker, 1984) in assuming that once a speaker (L1) is exposed to input, he or she can begin to limit his or her initial state (UG)—where all linguistic possibilities have the same activation potential—to build a network of language-specific lexical entries. In building a lexicon, therefore, children must be able to first identify that there is a certain concept in the world to express linguistically. In general, evidence from child acquisition points to the fact that children begin to build their lexical entries by first figuring out the morphophonological characteristics of the particular way to express a concept, and then figuring out the grammatical features required to combine those linguistic referents together. As such, children typically encode more obvious semantic concepts first, such as people, things, actions, directions and deixis (up, there, etc.) before they begin to linguistically encode more abstract semantic concepts such as number, tense, aspect, specificity, etc. Nevertheless, once a child begins to combine units of information (i.e. morphemes), he or she must know, to a certain degree, how those units come together to have combinatorial meaning from a language-specific point of view (i.e. what is interpretable for that speaker). As such, a child will slowly figure out from his or her input which
functionally-relevant semantic concepts are necessary to encode linguistically, and how those concepts are morphologically expressed, in order to build an interpretable linguistic expression in his or her particular language. The child will, thus, limit his or her universal language knowledge to converge upon a more restricted set of linguistic features encoded to particular lexical entries. These lexical entries vary cross-linguistically (grammatically-speaking), in terms of the formal semantic information that they are specified to encode, as well as in terms of their syntactic features that will allow them to combine with other lexical entries (which must be computable from a universal point of view). This internal structure of the grammatical part of the lexical entry contains all the formal features stipulated in a model of generative morphosyntax (i.e. what features are instantiated, if they’re strong or weak, interpretable or uninterpretable, etc.)

For example, for the expression Hasan saw a book (example from Jung Song, 2001, p. 162), the conceptual plan will specify what role the referent related to ‘book’ has in the event (theme, the thing ‘seen’), and what its semantic characteristics are in terms of definiteness, referentiality, number, etc. Depending on the language, these semantic features may or may not be encoded explicitly, and when encoded explicitly, may be encoded in different morphological ways. For example, in English, only definiteness is encoded linguistically here; definite arguments are implicitly referential, while indefinite arguments are ambiguous, as illustrated in (51) and (52).

(51) Hasan saw the book. (definite article, implicitly specific)
(52) Hasan saw a book. (indefinite article, ambiguous for specificity)

In English, in order to determine whether an indefinite argument is referential or not, the speaker and interlocutor must utilize discourse context to recover this information. In Farsi, however, both definiteness and referentiality are encoded explicitly; the referential direct object argument

---

29 The morphophonological information specified to the lexical entries will contain all the necessary language-specific information regarding the relevant aspects of sound and form in a language: phonotactics, syllabification, pitch and tone, etc.

30 Here, I use Farsi, not Spanish, to illustrate how different languages diverge with respect to their grammatical encoding of semantic concepts because this example is one of the most transparent that I’ve seen related to this type (semantics-syntax) of cross-language variability.
is marked for accusative case and remains unmarked for definiteness when the direct object is a
definite, referential argument.

(53) Hasan ketab-\textit{ra} did
    Hasan book-\textit{ACC} saw
    “Hasan saw \textit{the} book”

Nevertheless, when the speaker wants to indicate that the direct object is referential but not
definite, he or she must add the functional morpheme, \textit{yek}, to indicate the lack of definiteness for
the particular communicative context.

(54) Hasan \underline{\textit{yek}} ketab-\textit{ra} did
    Hasan \underline{\textit{a}} book-\textit{ACC} saw
    “Hasan saw \textit{a} book” \hspace{1cm} (a specific book)

Finally, when the speaker wants to indicate that the book is not referential, he or she only uses the
indefinite morpheme, but not the accusative case marking:

(55) Hasan \underline{\textit{yek}} ketab did
    Hasan \underline{\textit{a}} book saw
    “Hasan saw \textit{a} book” \hspace{1cm} (any book, non-specific)

The difference between Farsi and English in the expression of the direct object, therefore, is
related to the linguistic encoding of two semantic notions, definiteness and referentiality. While
referentiality and definiteness are linguistically encoded in two different ways in Farsi
(accusativity marking and indefinite morpheme), referentiality is not marked overtly in English,
but rather recovered from the discourse context. This example clearly shows how a basic
difference between languages lies in the morphosyntactic specification of certain semantic
features.

Apart from morphophonological differences, language differences (or similarities) arise,
therefore, due to the idiosyncratic specification of:

i) semantic concepts that are encoded linguistically in any one language;

ii) the way those semantic concepts are encoded linguistically (the internal structure of
the syntactic specification (the lemma)); and

iii) the way that the syntactic features interact to build a linguistic expression.

Child bilinguals, in the most idealized way, would build two independent monolingual-like sets
of lexical entries on the same developmental timetable as monolingual children, and subsequently
use them throughout their life in the most monolingual-like way. This idealized version of
bilingualism, however, is not a linguistic reality,\textsuperscript{31} and as such, in describing the bilingual lexicon, we must be able to account for how the lexical information interacts cross-linguistically and can be used by a bilingual in the formulation of a linguistic expression.

In terms of the general organization of the linguistic vocabulary (i.e. lexical and functional categories) of bilinguals, psycholinguistic research has provided strong evidence that one lexicon is responsible for storing all of the lexical entries of a speaker, regardless of whether they are mono- or multilingual (see review of early studies in Dijkstra, 2003 and Costa, 2006). The notion of non-selective lexical access is supported in empirical research examining both lexical and syntactic effects through priming and lexical/structural decision tasks. For example, Dijsktra, Grainger & van Heuven (1999) tested the reaction times of Dutch-English bilingual adults on lexical decision tasks and observed evidence of cross-language competition between words that were similar in form and meaning.\textsuperscript{32} Additionally, bilinguals accessed high frequency items faster than low frequency items, suggesting a role for frequency effects in the eventual selection of cross-language resources. Additionally, Meijer & Fox Tree (2003) and Salamoura & Williams (2006) showed evidence of non-selective access of syntactic features in cross-language priming tasks. The latter authors suggest that a lexicalist proposal, such as the one in Pickering and Branigan (1998, cf. Salamoura & Williams 2006), where syntactic priming is accounted for by “feature-based activation at the lemma level” (p. 299), can best account for the evidence of syntactic priming in bilingual speakers.

If bilinguals have only one lexicon, and it operates according to a lexicalist (and activation spreading) framework, then it is necessary to determine how bilingual speakers select the most language-specific features possible while limiting the random use of cross-language linguistic information in single-language utterances. In the next section, I will discuss the notion of language selectivity in bilingualism and what factors are associated with the differential

\textsuperscript{31} We are reminded here of Grosjean’s (1989) famous title: The bilingual is not two monolinguals in one person.

\textsuperscript{32} In psycholinguistic studies testing monolingual speakers, researchers have also shown that priming effects are not necessarily related to thematic similarities, but can occur for sentences that are congruent in surface structure only (i.e. \textit{by}-phrase in a passive can prime a \textit{by}-phrase that is the instrument of an action, and vice versa, Bock & Loebell 1990).
abilities that bilinguals appear to have in using strict language-specific resources in the analysis or construction of a linguistic expression.

### 3.3.1.2 Language Selectivity in Bilingualism

In the case of language selectivity, it was alluded to in Chapter 2 that the type of inhibitory mechanism available to bilingual speakers is unique to dual language contexts, in the sense that it is a mechanism that operates automatically to filter out linguistic information of the non-target language in contexts of communication where single-language utterances are most appropriate. As it turns out, however, this inhibitory mechanism may be unique to contexts of multilingualism, but is not necessarily unique to language. Psychological research on child and adult bilinguals has generally shown that the type of inhibitory control bilinguals have to filter out irrelevant linguistic information may be related to the functioning of executive control, a set of general cognitive processes operating across different cognitive domains. Research conducted using a variety of linguistic and non-linguistic methods (i.e. grammaticality judgments of ungrammatical and semantically anomalous sentences, dichotic listening task, Simon task, change-card sort task, among others) by Bialystok and colleagues (review in Bialystok, 2009) and Soveri et al. (2011) has shown that both child and adult bilinguals exhibit a cognitive advantage over monolinguals in the functioning of executive control. This evidence suggests that bilinguals have better-developed (automatic) control over this cognitive process compared to monolinguals, implying that filtering out irrelevant information is a process used frequently in contexts of dual-language exposure, and less in contexts of single-language exposure. Bialystok (2009) notes:

“[early child bilinguals] performed better than the other two groups [monolinguals and late child L2] on precisely the tasks that presented conflict for competing options that needed to be resolved for a correct response. This parallels the situation in which two competing language systems create a conflict for selection in bilingual speech production.” (p. 5)

---

33 Hernández and Meschyan (2006) also interpret differential fMRI readings for monolinguals and late adult bilinguals (on a picture naming task) to suggest that the function of executive control differs between these two groups.
If this executive control function is responsible for the filtering process of language information in bilingualism, and monolinguals do not exhibit as good control over this function, then it appears as though this type of cross-language inhibitory process is different than the process used to make choices among different language registers (which monolinguals are very good at), as I suggested earlier.

If we assume that the lexicon operates on the basis of a system of spreading activation, as outlined in Chapter 2, then the way in which this inhibitory mechanism operates may be to lower (or eliminate) the activation potential of items recognized as being irrelevant for the particular communicative act underway. While inhibition control (IC) is generally thought to be operational in most contexts of bilingualism to suppress non-target language resources in comprehension and production (Green, 1986, 1998; Costa & Santesteban, 2004; Kroll et al., 2006; Costa et al., 2006; Schwieter & Sunderman, 2008), there is some evidence to suggest that the most highly proficient bilingual speakers may be able to select target-language resources (at all levels, conceptually, lexically, etc.) directly without the use of an attentional control mechanism. Costa & Santesteban (2004) tested the processing costs associated with language-switching in highly proficient bilinguals (Spanish-Catalan) and Spanish/Korean speakers of L2/L3 English, Catalan or Spanish. The authors showed that only the most highly proficient bilinguals did not exhibit asymmetrical switching costs across languages, and as such, the authors suggested that these highly proficient bilinguals were able to select language-specific resources without the recruitment of some inhibitory mechanism, unlike the less proficient L2 speakers.

Nevertheless, Kroll, Bobb & Wodniecka (2006) affirm the “fundamentally nonselective” nature of a bilingual’s language system, and assert the following:

“Although there are circumstances that allow bilinguals to plan spoken utterances exclusively in one language without the influence of the other language, those circumstances are the exception, not the rule, particularly when speaking the L2.” (p. 127)

Here, the authors suggest that a variety of factors affect how “deeply the system remains open to interaction”, such as proficiency, dominance, language mode, the demands of the communicative task etc. (p. 127-128). They also suggest that while the lexicon (and its relationship to the semantic representation) is most commonly thought to be the locus of interaction, the authors point out that morphophonological information can have an effect on language selection,
especially depending on the relative proficiency of the speaker. While this type of system—where a backward flow of information from phonology to the lexicon can affect lexical choices in production—is not compatible with a Levelt-type model where information flow only happens in one direction, we must recall that Levelt’s model was constructed on the basis of a monolingual system. Like Kroll et al. (2006) suggest, bilinguals may make use of the overall system (i.e. different loci of the system) as a way to efficiently figure out aspects of language that are particularly challenging (i.e. underrepresentation, ambiguous input, etc.), especially at early stages of acquisition.

Nevertheless, despite bilinguals having expanded cognitive resources that aid in language selectivity (i.e. executive control), most research suggests, as mentioned above, that the relative proficiency of the speaker affects the recruitment of cognitive and linguistic resources in the analysis or construction of a linguistic expression in bilingualism. For example, in studies examining lexical selection, early learners of an L2 tended to be more form-dependent (morphophonological information), whereas more proficient learners tended to be more meaning-dependent (conceptual-syntactic information) (Sunderman & Kroll, 2006). Relative proficiency, therefore, has often been proposed as a deterministic factor of language selectivity in bilingual populations (see Juffs, 2009; Francis, 2011). What proficiency means, however, largely depends on your linguistic point of view. Schwieter & Sunderman (2008) define higher proficiency levels in terms of the recruitment of more procedural (automatic) language processes, and less reliance on attentional control to inhibit cross-language activation. They suggest, as Costa & Santesteban (2004) demonstrated, that bilinguals at the highest levels of proficiency do not typically need to use an inhibitory mechanism to suppress non-target resources because language-specific processing occurs automatically once language choice has been indicated as part of the conceptual plan. As such, Schwieter & Sunderman (2008) propose the Selection by Proficiency Model to account for the fact that the use of a general cognitive function, such as executive control, may be needed less to suppress the activation of cross-language resources as the bilingual gains more experience with (both) language(s). However, the authors also point out that in order to determine a certain threshold of proficiency for automatic processing, there must be some linguistic measure that correlates to lower processing costs in production. They note the need to tease apart aspects of a speaker’s language knowledge that lead to higher proficiency, and empirically demonstrate that lexical robustness may be one of the factors that allows a bilingual
to rely less on attentional control in the processing of one language. Here, lexical robustness is defined as “the familiarity with and frequency of access that leads to greater automaticity of retrieval of lexical items” (p. 216). According to the authors, this does not only mean relative vocabulary size, but also how strong those representations are, which they measure in a variety of on-line and off-line tasks. The relative strength of lexical representations, and the robustness of the system, lies on a continuum: the stronger the representation (enhanced through repeated access and retrieval), the more robust the system, and the less need the speaker has to rely on an inhibition mechanism for language selection.

Being a highly-proficient bilingual speaker, however, does not mean that attentional resources (i.e. executive control) will not be used in the processing of language, just that the speaker may rely less on these cognitive resources for functionally-motivated purposes. In Costa et al. (2006), highly proficient Spanish-English bilinguals employed both inhibitory control and direct automatized language-specific selection in language processing. The authors discuss differential uses of processing strategies for language selectivity in bilingualism according to factors such as language-pairings and the age of first exposure to bilingualism. Overall, therefore, while more highly proficient bilinguals may rely on a general cognitive function (i.e. executive control) less than unbalanced or less proficient bilinguals in order to select strict target-language resources, some general cognitive function associated with attentional control appears to be operational for linguistic purposes for all types of bilinguals. Older speakers, therefore, appear to have two jobs in acquiring a second language: they must learn the new language and learn how to recruit the appropriate functions of executive control for linguistic purposes. This latter feature of bilingualism may make late acquisition more challenging for an older speaker who has previously not used this cognitive function for linguistic purposes in any significant way.

3.3.1.3 A Language Model for Bilingual Representation and Use

In the current study, I assume that the system which underlies the production or comprehension of a linguistic expression for bilinguals and monolinguals is similar, except for two main differences:

1) when a speaker is bilingual, he or she has a lexicon that includes lexical entries from both languages for the same (or similar) concepts, and as such, can use cross-language resources for the grammatical formulation of an interpretable linguistic expression (for
2) when a speaker is bilingual, he or she will use the function of executive control for linguistic purposes, to inhibit the availability of cross-language resources when functionally or contextually appropriate.

Once a conceptual plan is formed, the lexicon is activated, and as such, the entire network of lexical entries is available for the formulation of a linguistic expression; the lexical entries that can encode the formal semantic values instantiated in the conceptual plan would have the highest levels of activation (in principle, in both languages). Nevertheless, declarative knowledge (involved in the conceptualization of the message) would feed the system so that the executive control function could operate to reduce the activation potential of the parts of the network that are irrelevant to the particular cognitive act (the non-target-language). In this way, in order to produce a strict monolingual-like expression, bilinguals must have a maximally effective executive control function that eliminates all cross-language information, and as such, the processing of language for bilinguals may require more cognitive effort than for monolinguals. Nevertheless, in some cases, bilinguals may have trouble in selecting a lexical entry in the target language to encode a particular semantic value of the conceptual plan because that lexical entry i) does not have a complete set of grammatical features (representational issue), or ii) can not be accessed efficiently (processing issue). In this way, the executive control function is overridden by a need the system has to search for information that can resolve the issue. Here, the system may diverge on the basis of age-related factors. For late bilinguals, the system will search the lexicon to resolve the problem, and as such, use cross-language resources more readily, especially at early stages of acquisition. Where the speaker finally comes to recognize that the L1 resources can not account for the particular property in the L2, the system may search UG, and as

Evidence from the work conducted by a variety of researchers (reviewed in Bialystok, 2009) suggest that this is so: bilinguals, as a group, are slightly slower at lexical access and retrieval than monolinguals. Bialystok (2009) notes:

“[Some] views attribute the reduction in lexical access to the conflict that is created by the competition from the corresponding item in the non-target language (Green, 1998). This competition requires a mechanism for controlling attention to the target language, possibly by inhibiting the interfering option. Generally, such conflict is resolved by the executive processes for control, attention, and switching. If these processes are involved in ordinary language production for bilinguals, then it is possible that their constant use in an ordinary and frequent context will have the consequence of transforming those processes through practice, making them more efficient and more available for a variety of applications.” (p. 5)
such, have access to properties that are not instantiated in the L1 lexicon. For early bilinguals, the system will search UG to resolve the problem, or where bilinguals have already acquired a strong representation of lexical entries for a particular grammatical domain and/or for one language (over the other), the system may also use cross-language resources.

Similarly, in cases where highly-fluent bilinguals choose to communicate in a way that is appropriate to their speech group (i.e. code-switching, etc.), the declarative information feeding the conceptualization phase will specify that the executive control function that filters cross-language resources is not needed at any particular moment of a speech act. In cases of code-switching, for example, the bilingual does not need to filter out irrelevant cross-language information, but rather has developed a system that is strong enough where procedural linguistic knowledge can determine which cross-language lexical entries and features are compatible cross-linguistically. As a result, the linguistic expressions formulated are still interpretable from both a universal (computational) and language-specific (PF/LF) point of view. Also, in the case where bilinguals are able to communicate in a more bilingual mode of communication, they may do so automatically because it is more economical from a cognitive point of view: the executive control function does not have to operate as strongly to keep a strict division between language resources, as is required in a monolingual context of communication, and as such the bilingual exerts less cognitive effort to formulate a single-language utterance that is interpretable in a bilingual context. Finally, over time a bilingual may need to rely less and less on the function of executive control to maintain maximal differentiation between languages, especially for the most frequently used items/structures (recall frequency effects of Dijkstra, Grainger & van Heuven, 1999) Nevertheless, given the robust evidence that there is for cross-language effects at even the most advanced levels bilingualism, it appears as though the bilingual relies on the function of executive control to some degree through the lifespan as a speaker’s cognition and exposure to language change over time.

So, if the process by which bilinguals filter out cross-language information in single-language utterances is a function of executive control, and it is typically better-developed in bilinguals, then a question arises regarding age-related differences: How hard is it for an early child bilingual to use this executive control function during grammatical development, and is an older child or adult able to recruit this cognitive function for language purposes when they have
always been a monolingual? In the first case, most evidence on early child bilinguals shows that they are able to make appropriate linguistic choices from as early as 2;0 given different communicative variables, including the linguistic abilities of their interlocutor, the language environment, etc. (Genesee, Nicoladis & Paradis, 1995; Genesee, Boivin & Nicoladis, 1996; Lanza 1997). From the earliest stages of grammatical development, therefore, the executive control function of bilingual children appears to be operational and functioning in an efficient way. This is consistent with most of the evidence of grammatical development in child bilingualism, where bilingual children (who are not strongly dominant in one language) rarely exhibit qualitative bilingual effects between 2;0 and 4;0 (see review in Paradis, 2009). Most of the developmental language of early bilinguals converges to that of monolinguals of the same language, they simply may exhibit different rates of development due to the different patterns of exposure to input that they experience. Nevertheless, as mentioned above, some bilinguals who are dominant in one language (linguistically) may combine their language resources in order to functionally communicate in their weaker language. In this case, dominant bilinguals may not have enough knowledge of their weaker language to make judgments regarding the interpretability of a badly-formed utterance when using cross-language resources.

Using cross-language resources appears to be especially characteristic of late child and adult L2 acquisition, especially at the early stages of acquisition when new bilingual speakers are heavily L1 dominant. This observation is captured in the proposals of L2 acquisition which suggest that L1 transfer constitutes the initial state of the L2 grammar for a late learner (i.e. Schwartz & Sprouse, 1996). The acquisition task for these speakers, therefore, is to not only learn the language-specific features of their L2, but also to learn how to use executive control functions for linguistic purposes; they must learn how to efficiently eliminate the (strong) activation potential of lexical entries in their L1 once their L2 lexical entries have been established. As such, the efficient operation of executive control in late acquisition is largely determined by the degree to which a speaker can build language-specific lexical entries in the L2 and recruit the efficient use of the function of executive control to filter out the strong linguistic resources of the L1. The apparent ease that children have compared to adults in the acquisition of an L2, therefore, may be related to the differential ways in which children and adults are able to overcome their strong reliance on L1 resources to formulate L2 linguistic expressions, and as such, recruit executive control functions in a maximally effective way. As such, factors such as
language-confidence, anxiety, etc. may also affect the degree to which the executive control function may operate for an adult who has never before used it for language purposes. Nevertheless, since there is evidence of late bilinguals who do acquire near-native grammars, and given that executive control is a general cognitive function of all normally-developed humans, then there is no principled reason to suggest that the function of executive control in language can not operate in a maximally efficient way for a late language learner; it just happens that it often does not because of a variety of language-internal (representational/processing issues) or language-external (relative language dominance, language anxiety) factors. Further, as mentioned previously, many bilinguals may choose to communicate in a way where cross-language resources are used to construct a linguistic expression in a communicative interaction with another bilingual (code-switching, bilingual mode of communication, etc.). Consequently, the functioning of the executive control process responsible for inhibiting cross-language resources appears to operate variably according to the relative linguistic proficiency and/or communicative needs of the speaker. In the case where the speaker exhibits functional or contextual cross-language effects, the activation potential of the cross-language resources may remain high enough so that the speaker can rely on residual activation effects in the selection of a target language feature (quantitative effects), or so that the speaker can directly use linguistic information from the non-target-language in the formulation of a linguistic expression in the target language (quantitative or qualitative effects).

Here, we are introduced to the two types of (single-language) interaction effects commonly cited in studies on bilingualism: cross-language influence and transfer. In the current study, cross-language influence refers to

i) the residual activation of cross-language resources that leads to a heightened activation potential for a lexical entry (grammatical and/or morphophonological information) in the target language; or

ii) an indirect influence of an invariable word order/use in the non-target language on word order variation in the target language during morphophonological encoding.

The non-target-language, therefore, provides an indirect source of influence that results in a (unidirectional) quantitative bilingual effect in the target language. Transfer, on the other hand, is a process by which morphosyntactic (and/or morphophonological) information is recruited directly from the non-target-language for the formulation of the linguistic expression in the target
language; in this way, either a quantitative effect results (with more strong interpretational consequences), or a qualitative effect results, where it may appear that the speaker is lexicalizing the syntactic frame of the non-target-language with morphophonological information from the target language. I will systematically describe these two types of cross-language interaction effects below in section 3.3.3, along with how these effects operate alongside a general bilingual effect.

In sum, the language system outlined in Chapter 2 functions in relatively the same way for bilinguals as monolinguals. Where these two populations diverge, however, is in the availability of cross-language resources for the formulation of a linguistic expression according to contextually- or functionally-relevant needs in contexts of bilingualism, and in the use of certain functions of executive control for linguistic purposes in the suppression of cross-language information. In figure 3.1, I have constructed this model in a way that represents different phases of formulation of a linguistic expression, as well as the areas of the system responsible for the procedural and declarative information associated to variability in the ways that any particular speaker will formulate the same desired meaning. The exact way that the system functions for bilingual speakers also varies in terms of a variety of language-internal factors (language typologies) and language-external factors (age of exposure to bilingualism, language dominance, etc.).

Figure 3.1. A Language Model for Bilingual Speakers
In general, three types of (single-language) bilingual language behaviour are possible with such a system: a general bilingual effect, cross-language influence and transfer. Below, I will consider these three in more systematic detail along with empirical evidence from studies on child and adult bilingualism of the ways that child and adult bilinguals have exhibited use of these three types of bilingual language behaviour in early and mature stages of development.

3.3.2 A Typology of Bilingual Language Behaviour (Single-Language)

Since some of the earliest studies on L2 acquisition, the influence of the L1 on the development and outcome of exposure to the L2 has played an important role in constructing theories related to the initial state and what those acquiring an L2 are able to ultimately attain (Krashen, 1982; Bley-Vroman, 1989; Selinker & Lakshmanan, 1993; Birdsong, 1992; Lardiere, 2000; White, 2003; Montrul, 2008). Similarly, most theoretical views of early bilingual development incorporate some notion of language influence (or interference) to account for evidence of independent, but selectively non-autonomous grammars through the course of bilingual development (Genesee, 1989; Meisel, 1989; Paradis & Genesee, 1996). In many studies examining early and late bilingual development the terms transfer and cross-language influence (or interference) are often used interchangeably (Odlin, 1989; Müller, 1998; Ionin & Zubizarreta, 2010). Some attempts have been made to clear up this terminological confusion, relating transfer to a more obvious and direct influence of the one grammar (i.e. L1) on the other grammar (i.e. L2), while cross-language influence being related to some sort of (indirect) quantitative bilingual effect that occurs uni- or bi-directionally due to the bilingual nature of the individual’s exposure to input (Müller & Hulk, 2001; Pérez-Leroux, Pirvulescu & Roberge, 2009).

Nevertheless, if we adopt an approach where all the basic components of the language faculty are innate and available on a continuous basis from child- to adulthood, then the types of bilingual language behaviour exhibited across the lifespan should be continuous as well. Whether a particular speaker (child or adult, native or non-native) exhibits the use of one type or another will have to do with the way that the different components of the system function and interact amongst each other and in relation to the function of general cognition based on the experiential- and cognitive-related factors associated with age of first exposure to bilingualism. While empirical evidence of language development from the initial stage to mature grammars in early and late bilingualism will be reviewed from a chronological perspective in Chapter 4, this section
will systematically describe how the notions of language transfer, cross-language influence and a general bilingual effect are accounted for in the current study given the language model presented above.

### 3.3.2.1 A Systematic View of Quantitative and Qualitative Effects

As mentioned previously, a bilingual speaker may produce a range of bilingual-like utterances that diverge from those of monolinguals in very subtle ways (i.e. quantitatively) or in more obvious ways (i.e. qualitatively), and somewhere in between. Before considering this range of behaviour, however, let us review the ways in which cross-language variation can be accounted for given the language model used in the current study. First, each language is unique in its particular linguistic encoding of many of the semantic features expressed in a conceptual plan. Earlier we saw examples of how Farsi encodes different morphemes to encode the notions of definiteness and referentiality to an accusative argument while English only encodes the notion of definiteness explicitly.

A second way that languages may differ is in the linearization and/or use of certain sentential constituents/items. From a generative perspective, basic linearization (word order) arises from whether substantive categories are left- or right-headed. For example, English is a VO language, and as such, verbal constituents are left-headed; conversely, Chinese is an OV language, and as such its verbal constituents are right-headed. When languages exhibit categorical restrictions on the headedness of constituents (verbal, prepositional, nominal), bilinguals rarely exhibit cross-language interaction effects once they have basic knowledge of the language(s) being acquired. However, the example of a right-headed relative clause uttered by a Chinese-English bilingual child, repeated below in (56), shows that these strong qualitative effects do occur and as such, are cases where speakers appear to draw upon linguistic information (grammatical and morphophonological) from the lexical entries of both languages for the formulation of a linguistic expression.

(56) Where’s the [Santa Claus give me the gun]? Where’s [the gun Santa Claus gave me]?

More common qualitative effects, however, are when speakers linearize constituents (not head-related) or express morphosyntactic elements not instantiated in the target grammar. These types of qualitative effects are often obvious influences of the non-target grammar (often called
negative transfer), but are used in ways that typically do not affect the gross interpretability of the utterance. For example, in producing English relative clauses, Arabic-English bilinguals may often use a resumptive pronoun in the relativized clause, an obvious qualitative bilingual effect that researchers have suggested occurs based on negative transfer from Arabic (Kharma, 1987; Kassabgy & Kamel Hassan, 2000):

(57) This is the house which I live in it.

Given the evidence that there is for a range of (strong to mild) qualitative effects in both child and adult bilingualism (especially in early stages of late child and adult L2 acquisition), I will assume here that while a complete lexical entry is activated (and can be selected), the grammatical information (Levelt’s lemma) and morphophonological information contained as part of the lexical entry are relatively independent in the sense that the grammatical information from one language can be used with the morphophonological information of the other in the formulation of a linguistic expression (see Separation Hypothesis of Beard, 1987, cited in Spencer, 1991, Lardiere, 2000), among others. In this way, we may be able to account for cases where it appears as though a speaker is using a syntactic frame from the non-target language with morphophonological forms (words) from the target language.

Nevertheless, the more common type of bilingual language effects (especially for proficient speakers) occurs when the target language exhibits word order variability that the target-language does not. In these cases, the resultant bilingual utterances may be related to the linearization and/or use/omission of sentential constituents/items in a way that diverges from that of monolingual speech patterns. Two areas of the language system may be responsible for this type of linguistic variability: 1) the activation and selection of lexical entries (pre-Spell Out); or 2) the linearization and/or insertion of morphology (post-Spell Out). A general bilingual effect and cross-language influence are described together below because both are effects that can occur due to the selection of language-specific lexical entries in the target language prior to grammatical encoding or due to the linearization/morphophonological processes that take place after a linguistic expression has been encoded grammatically. Further, these types of bilingual effects are systematically similar in the following two ways:
1) the resultant bilingual language behaviour is congruent to forms/structures used by monolinguals but used in a quantitatively different way at a given point of development, or for a particular context of use; and
2) only target-language resources are used by the bilingual in the formulation of their linguistic expression; the apparent divergence results from using universal/unmarked features more than monolinguals or due to an indirect effect provided by the residual activation of lexical entries or linearization patterns of the non-target language.

In the next section, I will describe in more detail how these two types of quantitative effects arise in the language of bilinguals, and the proposals that have been made to account for evidence of this type of variability in child and adult bilingualism.

### 3.3.2.2 General Bilingual Effect and Cross-Language Influence

Pérez-Leroux, Pirvulescu & Roberge (2009) proposed the Default Retention Hypothesis to account for contexts in which early bilinguals exhibit a persistent use of universal linguistic options (in one or both languages) for the expression of a particular grammatical domain past the point at which monolingual children of the same language have converged to the language-specific features for the same domain. In a series of work (2006; 2009; 2011; Belzil et al, 2007) examining the use of pronominal objects by English monolinguals, French monolinguals and French-English bilinguals between the ages of 3;0 and 7;0, the authors observed that French-English bilinguals used higher rates of illicit object omissions in both languages past the stage at which monolinguals of both languages had converged to target-like grammars for this grammatical domain, illustrated in (58) to (60).

(58) She’s drinking **it**. [the milk]
(59) **Elle** boit. [**le lait**]
(60) i. She’s drinking **Ø**. [the milk]
    ii. **Elle** **Ø** boit. [**le lait**]

The authors, thus, proposed that the bilinguals retain the universal option available for transitive contexts—a null referential object—longer than monolinguals because they require more relative time with input to converge to the language-specific grammar for each language: phrasal pronouns in English, and clitics in French.
This type of bilingual effect is defined in the current study as a general bilingual effect, and is commonly observed in the grammars of early bilinguals who require more relative time (chronologically) than monolinguals to converge to the language-specific features for a particular grammatical domain in one or both languages (Pérez-Leroux, Pirvulescu & Roberge, 2009; Gu, 2010). As such, here the apparent effect can emerge in only one grammar (appearing as a so-called unidirectional effect), or can emerge in both grammars, leading to what has often been called a bidirectional effect. Whether the effect is unidirectional or bidirectional often has to do with the relative complexity of the particular grammatical domain in each language, and as such, the relative threshold of input required to converge to the most language-specific features for that grammatical domain. Nevertheless, since this type of bilingual effect does not proceed on the basis of the influence of one grammar to the other, but rather in the persistent use of (universal) developmental features (common to all speakers of the language), the descriptive terms unidirectional and bidirectional are not necessarily appropriate. Nevertheless, to be most consistent with the terminology employed in the field, the terms unidirectional and bidirectional will be used here to refer to contexts where one or both languages of the bilingual exhibit bilingual language effects in development or at mature stages, and should not be understood as terms that relate to the direct (or mutual) effect of one language on another.

While the Default Retention Hypothesis best applies to contexts of early bilingualism where the child selects universal (often persistent developmental) features over language-specific ones (pre-Spell Out), there is also evidence from contexts of L2 acquisition where older child and adult bilinguals resort to unmarked features of the target language to express a particular semantic concept that is not necessarily encoded in the non-target language (their L1). In these cases, the speaker does not fail to provide a morphological expression of the particular feature, but rather defaults to the unmarked form available for that particular expression. For example, in cases where a speaker does not encode grammatical gender to nominal constituents in their L1, they may have trouble in converging to the most language-specific system where gender assignment and agreement are characteristic of their L2. In empirical work on Spanish-English bilingualism (adult L2 Spanish), White et al. (2004) and McCarthy (2008) showed that English L2 speakers of Spanish may not have difficulty in acquiring the features responsible for the expression of nominal gender in Spanish, but may have difficulty in expressing the most target-like instantiation of that feature for a variety of gender agreement contexts in production and
comprehension. McCarthy (2008) proposed the Morphological Underspecification Hypothesis to account for contexts in which speakers use a systematized strategy of underspecification in the expression of tense and gender in Spanish. She showed that the errors made by intermediate-advanced L2 Spanish speakers for tense and gender agreement corresponded to the unmarked morphological forms for verbal and nominal inflection: 3rd person singular and masculine singular morphology, respectively. Here, the resources of the non-target language do not affect the way that the bilingual speaker expresses tense and gender morphology; rather, the bilingual effect is due to the overgeneralization of unmarked target language resources. As such, the resultant error may occur because the speaker lacks grammatical knowledge of gender-related syntax or has trouble in expressing the features morphologically. In the first case, the speaker may have a functional need to find a way to express a determiner or adjective when he or she lacks the knowledge of the particular gender of a nominal referent, or lacks knowledge of the features that lead to gender agreement. Consequently, the resultant bilingual expression is representationally different than the one expressed by monolinguals. In the second case, the speaker may have knowledge of the relevant features associated to the expression of target-like gender (i.e. assignment and agreement), but resorts to unmarked features in the expression of the utterance for economy purposes: either the speaker needs to sacrifice a part of his linguistic expression because of the cognitive burden involved in producing a target-like utterance, or the speaker has identified a context in which he can shift into a more economical bilingual mode. In this case, the bilingual expression is representationally congruent to that of a monolingual, but morphophonologically expressed in an unmarked way. This process is different than that of a mapping problem, as stipulated in the Missing Surface Inflection Hypothesis (Haznedar & Schwartz, 1997; Prévost & White, 2000), which I will consider in more detail below.

In sum, child and adult bilinguals may use universal or unmarked resources to produce a bilingual utterance because of a functionally- or contextually-motivated need to express themselves (in the most efficient way possible). The resultant general bilingual effect is not due to any direct or indirect influence of cross-language resources, but rather arises due to the individual’s unique cognitive abilities (linguistic knowledge and/or procedural capabilities) given his or her experience with bilingualism at any particular point in development. As such, this type of general bilingual effect can arise in one or both languages of the bilingual, depending on the complexity of the language domain and/or the bilingual’s relative experience with input for any
one language (i.e. age of first exposure, relative quantity of exposure, etc.). Pérez-Leroux, Pirvulescu & Roberge’s (2009) Default Retention Hypothesis and McCarthy’s (2008) Morphological Underspecification Hypothesis appear to be able to account for the evidence from bilingual acquisition contexts that both child and adult bilinguals may resort to universal and/or unmarked features for one or both languages depending on a representational or processing need the bilingual has to express a concept in the most simple and efficient way possible.

Once a child or adult bilingual has begun to acquire some knowledge of language-specific lexical entries for a particular grammatical domain, he or she may come to realize through his or her exposure to surface syntax that there is more than one linguistic way to express a number of semantic features associated to that domain in one language, but only one invariable way to express that group of semantic features in the other language. For example, in Spanish, there are two ways to express pronominal subjects: with null pronouns in unmarked discourse contexts and with overt subjects in marked discourse contexts (to focus the subject constituent). In principle, all speakers would start out with a universal null subject applying to all discourse contexts, and then with input from the target language, they would begin to acquire more language-specific features that, in effect, would limit the use of a more language-specific null subject. In the case of English, the speaker comes to know that the language is a non-null subject language, and as such, the use of a null subject is (mostly) categorically restricted. As such, the semantic features instantiated for the basic morphosyntactic features of the external argument (person, number, case) are invariably expressed on the overt subject; the discourse-related feature of subject focus is relevant at the morphophonological level of analysis where an English speaker applies contrastive stress to the overt subject constituent. However, in the case of Spanish, both basic morphosyntactic features and discourse-related features apply at the grammatical level of analysis. With exposure to (monolingual-like) Spanish input, children eventually come to know that the use of the overt subject is restricted to marked discourse contexts, while the null subject is used in unmarked discourse contexts; the basic morphosyntactic features of the external argument are recovered on the verbal morphology of the tensed predicate. Even if a bilingual comes to acquire language-specific lexical entries for this grammatical domain, the relative levels of activation for the competing variants in Spanish may be different than that of a monolingual because of the associated (strong) activation of the overt subject in English. As such, even if the conceptual plan stipulates that the resultant linguistic expression is to be interpreted according to
unmarked discourse-related features, the residual activation from the English lexical entry may enhance the activation level of the lexical entry for the Spanish overt subject, thereby sending the overt subject through to be grammatically encoded even though the interpretation is unmarked from a discourse point of view. Instead of resorting to the unmarked feature, therefore, the bilingual uses the more marked feature because of an indirect cross-language influence. Here, we can discuss this type of influence in terms of directionality (unidirectional) because the influence will proceed from the language with less variability to more variability for a particular grammatical domain (as long as the surface syntax of the two languages is congruent). Further, we expect to see only quantitative effects due to cross-language influence; that is to say, more or less of a certain target-like grammatical expression is used by a bilingual given a particular communicative/discourse context.

Similar to a general bilingual effect, cross-language influence can occur because of a lack of knowledge on the part of the bilingual of the language-specific features that regulate the use of two competing grammatical options (representational issue) or may be due to a cross-language effect that occurs at the level of processing. In the latter case, there are two areas in the language system where processing may affect the eventual bilingual utterance. First, as mentioned, the residual activation of the invariable lexical entry from English may enhance the activation level of the overt subject in Spanish, leading to its selection for an unmarked discourse context. Here, all language-specific components of the system are operating in a target-like way, it is the function of executive control that may fail to apply in a maximally effective way to inhibit the activation of the lexical entry in English. This recruitment of executive control may be problematic for an adult bilingual because of his or her lack of experience in using this cognitive function for linguistic purposes up until adulthood, or may be relaxed when a speaker is in a bilingual mode of communication.

Further to pre-Spell Out effects, cross-language influence can have an effect at the stage of morphophonological encoding (after grammatical encoding has taken place). Prévost & White (2000) proposed the Missing Surface Inflection Hypothesis (MSIH) to account for evidence that bilingual speakers often leave off morphological inflection during production even when they show target-like knowledge of features that underlie a particular grammatical domain. While the MSIH may be viewed as a competing theory to that of the MUH proposed by McCarthy (2008), the variability of syntax-morphology mapping invoked by the MSIH may be able to account for
post-Spell Out cross-language effects that can’t be accounted for under the MUH (which proceeds as a general bilingual effect). For example, Lardiere (2000; 2006a) presented now-famous data from Patty, a Chinese speaker of L2 English who, despite exhibiting high proficiency on a variety of linguistic measures and target-like knowledge of tense in English, exhibited non-target-like behaviour in her use of tense morphology in production (optional use of past tense morphology in talking about past events):

(61) my mom also speak Mandarin
(62) so I wrote and speak fluently

This evidence of persistent, so-called residual optionality (Sorace, 1993) in the spontaneous production of Patty at relatively advanced levels of L2 proficiency led the author to propose that this particular speaker had experienced a fossilization in her ability to map the English past tense feature to the appropriate morphological forms in production, and instead rely on a more familiar (L1) approach, the recovery of time-related information through discourse context. Lardiere (2003) further proposed that indirect influence from the L1 can provide bilingual speakers with information regarding ways in which to linguistically express features of their L2. As such, cross-language influence in the post-Spell Out phase is differentiated from a general bilingual effect because the invariable patterns of morpheme use/linearization in the non-target language can affect the resultant use/linearization of morphemes in the target language. Lardiere concludes that “we should at least consider the possibility that differential tendencies among learners’ L1s to omit elements which are recoverable from the discourse context (or elsewhere in an utterance) might distinguish Chinese learners, from, say German learners of English”, who arguably omit tense marking in English much less at advanced stages of L2 proficiency (Lardiere, 2003, p. 188).

Further, cross-language influence in the post-Spell Out phase is different than that of the pre-Spell Out phase because the resultant expression is not badly-formed from a semantic-pragmatic point of view, but rather diverges (quantitatively) from the mature monolingual target in a way that is clearly associated to the articulation of morphology in the non-target language (i.e. patterns of inflection or linearization).

To conclude, I must acknowledge that it is often difficult to differentiate between a general bilingual effect and cross-language influence, especially for cases where the universal or unmarked form for a particular grammatical domain is linguistically congruent with the language-specific form in the non-target language. For example, if a child is acquiring French
and Chinese simultaneously and exhibits more object omissions in French than monolingual children at any particular stage of development (i.e. appears delayed for this grammatical domain), then it may be that the child is persistent in his or her use of universal options (a referential null object) for that grammatical domain because of his or her altered patterns of exposure to French (compared to a monolingual), or is using cross-language resources from Chinese (a referential null object) to resolve the learnability problem associated with acquiring the language-specific morphosyntax of French transitive sentences. In this way, for certain grammatical domains, it may be necessary to test different language pairings to determine if any quantitative effect associated with the use of universal or unmarked forms is the result of the use of cross-language resources or simply the result of the altered patterns of exposure that accompany contexts of bilingualism (or the degree to which each one has an effect). Of course, one is not necessarily exclusive of the other. As the process of acquisition occurs, speakers can rely on both cross-language resources and universal/unmarked resources to resolve any difficulty they have in the formulation of a target language utterance. As Lardiere (2003) pointed out, Chinese may enhance the tendency for an L2 English speaker to not use past tense morphology, but we must recall that using an unmarked verbal form in L2 English is certainly not unique to speakers of non-tense languages (Hawkins & Liszka, 2003).

Differentiating between a general bilingual effect and cross-language influence may end up being problematic in the consideration of the use of pronominal clitics in Spanish infinitival constructions by Spanish-English bilinguals. English speakers only use the unmarked way to formulate a complex transitive predicate ([aux/VP [VP-pron]]); consequently, if Spanish-English bilinguals exhibit higher rates of use of the unmarked variant (enclisis) in production, then two reasons for this quantitative difference are possible:

1) a general bilingual effect where the speaker uses the unmarked form as the structural default due to his or her relatively less experience, as a bilingual, in using the language-specific structure-building process of Spanish (verb incorporation);

2) cross-language influence where the indirect influence of English leads to the use of the unmarked variant in Spanish (due to residual activation effects of English lexical entries, or due to the indirect influence of the linearization of pronominal objects in English).

I will consider this issue more in Chapter 7 after presenting the data for the two groups of bilingual speakers examined in this study.
3.3.2.3 Transfer

In order to maximally differentiate between the indirect and direct influences that linguistic knowledge from the non-target language can have upon the grammatical formulation of a linguistic expression in the target language for a bilingual, here I systematically define transfer in terms of the use of linguistic resources from both the target and non-target language in the formulation of an utterance in the target language. I assume, therefore, that the two parts of a lexical entry—the syntactic and morphophonological information—are related, but relatively independent, as stipulated in the Separation Hypothesis (Beard, 1987). In the case of a bilingual speaker who has two cross-language lexical entries that are highly activated by the same semantic features, he or she can select the syntactic features of one while using the morphophonological information of the other (for certain items/constituents or for the whole utterance). This type of direct use of cross-language resources may occur when bilinguals lack target-language knowledge (representational issue) or have a lack of robust knowledge (L2 resources don’t win out strongly over L1, processing issue) of target-language features to formulate an interpretable linguistic expression, and thus, have a functional need to search the system for a way to resolve the representational or processing issue. Above, we saw that early child bilinguals tend to resort to universal features to resolve representational issues; however in contexts of L2 acquisition (especially for older children and adults), the L1 provides a more accessible source of linguistic information, and as such, especially at the early stages of acquisition, L2 speakers may exhibit transfer effects in their formulation of linguistic expressions in their L2. As mentioned, once L2 speakers begin to acquire language-specific knowledge in their L2, they must also learn how to recruit the function of executive control to inhibit the lexical entries of their L1 in order to formulate the most language-specific utterances in the target language. Consequently, even if L2 speakers become quite proficient in their knowledge of certain grammatical domains in the L2, they may still exhibit transfer effects in production because the functioning of executive control may not be maximally efficient due to a number of language-external factors (i.e. cognitive load, language anxiety, over-reliance on self-monitoring, meta-linguistic awareness, etc.).

While transfer effects are most common in the (early) grammars of L2 speakers, early child bilinguals have also been shown to exhibit this type of interaction effect. Most research points to two main contexts when early bilinguals employ grammatical transfer:
1) when the languages they are acquiring are asynchronous for the development of a particular grammatical domain;
2) when the child is linguistically dominant in the non-target language and resorts to cross-language resources (instead of universal resources) to formulate an expression in the target language.

As previously discussed, there is evidence from Chinese-English bilinguals and German-English bilinguals (among others) to suggest that early bilinguals use cross-language resources as a functional way to express more complex structures in the target language that have already been acquired in the non-target language in around the third year when grammatical development is occurring at a relatively rapid pace. This type of boosting effect for bilinguals is illustrated in (63)-(65).

(63) Where’s the Santa Claus give me the gun?
    “Where’s the gun that Santa Claus gave me?”
    (child 2;04; Yip & Matthews, 2000, p. 204)

(64) ich mochte tragen dich
    I want carry you
    Target: ich mochte dich tragen
    (child 3;02; Döpke 2000, p. 214)

(65) Simone is zu klein zu gehen in die kita
    Simone is too little to go to the day-care centre
    Target: Simone ist zu klein (um) in die kita zu gehen
    (child 2;07; Gawlitzek-Maiwald & Tracy, 1996, p. 918)

While Yip & Matthews point out that the gross violations of grammatical structure are allowed by this Chinese-English bilingual speaker because of her relative dominance in Chinese, Gawlitzek-Maiwald & Tracy indicate that the preferred language of the child in their study is German, and the most common language spoken to her is German. As such, it may not be the child’s relative patterns of dominance (related to uneven exposure) that leads her to use an English grammatical structure to express a German utterance, but rather the fact that she has already acquired a full representation for infinitival structures in English, but is still working through the acquisition process of acquiring the German rules for predicate structure (V2 grammar); that is to say, the input threshold to acquire English infinitival predicates is lower than that of German. This relatively earlier acquisition of predicate structure in English, therefore,
provides a bilingual child with resources that a German monolingual child doesn’t have at that same developmental moment to build infinitival predicates in German (i.e. English provides a structural hypothesis for infinitival structures). Given that this particular type of cross-language interaction proceeds on the basis of a functional need to formulate a grammatical expression for a particular conceptual plan, the types of mixed language utterances exhibited by early bilinguals may not necessarily conform to the systematic rules that more proficient bilinguals utilize in their use of code-switching (which is also a case of transfer—direct use of cross-language resources).

For example, Gawlitzek-Maiwald & Tracy (1996) showed that they observed a German-English bilingual using bound inflectional morphology (auxiliaries in German with ge- marking on past participles + English past participles) in a way that more proficient bilinguals may avoid when using cross-language resources for contextually-motivated communicative purposes (i.e. code-switching):

(66) ich hab geclimbed up. (child 2;04; p. 914)
    I have pp-climbed up

(67) ich hab gemade you much better. (child 2;04; p. 911)
    I have pp-made you much better

(68) du hast gebuyed them? (child 2;03; p. 911)
    you have pp-buyed them

Overall, bilingual speakers can use language transfer as a functional way to resolve representational deficits in the process of learning a new language or at mature stages of acquisition when they have indeterminate knowledge for a particular grammatical domain. Given that older children and adults start off as highly imbalanced bilinguals at the early stages of acquisition, it is not surprising that they exhibit high rates of qualitative effects in their L2 at this stage. Nevertheless, early child bilinguals may also exhibit transfer effects when the two languages they are acquiring exhibit asynchronous timetables of development in around the third year for a particular grammatical domain (they are linguistically dominant for that domain at that particular developmental moment) and/or when the child experiences patterns of exposure that highly favour one language over the other (overall linguistic dominance, but not necessarily for all domains). Further, in addition to acquiring the language-specific grammar for a particular grammatical domain, bilinguals must also be able to recruit the executive control function in a maximally efficient way so as to limit the influence of the non-target language in the grammatical formulation of a target language utterance. Here, age effects may be relevant, as it appears as
though adults learning a new language may have a harder time using this executive control function for linguistic purposes as efficiently as children, given the relative rate of qualitative effects that persist in the production (despite complete grammatical knowledge) of late (adult L2) bilinguals. I will address this issue further in Chapters 4 when I discuss the notions of ultimate attainment in contexts of L2 acquisition and how age-related cognitive and input variables may be able to explain the apparent age-related differences between child and adult bilinguals.

Finally, while the main difference between transfer and the other two types of bilingual effect described above (cross-language influence and a general bilingual effect) is that the bilingual produces grammatically divergent utterances (qualitatively) than are observed in the developmental or mature language patterns of monolingual speakers, the degree to which this type of transfer may persist at more mature stages of bilingualism depends largely on the speaker’s ability to recognize the grammatical violations of these structures and/or the relative degree of grammatical violation that such qualitative effects produce.

This persistent use of combined cross-language resources may eventually result in a stabilized bilingual representation, a phenomenon that has been shown to be common in language contact communities (Silva-Córvalan, 1994, 2003; Sánchez, 2004; Toribio, 2004). As such, when bilinguals (especially early bilinguals) instantiate features of a grammatical domain in a way that combines morphosyntactic information of both languages, but does so in a way that doesn’t produce gross grammatical violations (and as such, lack of intelligibility), this form may stabilize (depending on a variety of language-external factors) and possibly develop a grammatical status of its own for a particular community of speakers. I will consider these issues of bilingual variability and interpretability in more detail in the next chapter when I present an overview of contexts of dual language exposure from age-related and chronological perspectives with a focus on the issue of attainment in early and late bilingualism.

3.4 Chapter Summary
In this chapter I have considered how the language model outlined in Chapter 2 can be adapted for contexts of dual language exposure. To accommodate the unique cognitive needs of bilingual speakers, this bilingual language system differs in two fundamental ways: 1) the bilingual lexicon instantiates (competing) cross-language lexical items; and 2) the function of executive control is used for linguistic purposes as a way to inhibit the activation of non-target-language resources in
single-language utterances. In general the activation of lexical items is non-language-specific; as such in when bilinguals have stronger representations/access to certain lexical items in one language (in development or at mature linguistic stages), the resources of the non-target-language may contribute to the formulation of a linguistic expression, leading to what appears to be a bilingual interaction effect. Common types of interaction effects are cross-language influence and transfer. In the first case the residual activation of lexical items/features from the non-target-language enhances the activation of those in the target language; in this case only quantitative cross-language effects result. In the case of transfer, the linguistic resources of the non-target-language are directly selected for the expression of an utterance in the target language, and as such, qualitative or quantitative effects may emerge.

These interaction effects, along with the availability of unmarked/universal representations, can account for the types of bilingual language behaviour seen in developmental and mature language stages. For the current study, it is important to differentiate between a general language effect (greater reliance on unmarked structures) and cross-language influence (residual activation effects of the non-target-language) as evidence of the greater use of one structural variant over the other (proclisis vs. enclisis) may have implications for how we consider the processes of stabilization and attrition in early bilingualism and so-called complete acquisition in late bilingualism.

In the Chapter 4 I will consider these two contexts of dual-language exposure in more detail and consider the factors that have been proposed to account for differences in linguistic outcomes due to the age of exposure to bilingualism.
4 Early and Late Bilingual Development and Attainment

4.1 Introduction

The aim of this chapter is to review the processes of early child bilingual and late adult second language acquisition and to discuss the impact that the age of exposure to bilingualism has been shown to have on eventual linguistic outcomes in adulthood. Analyses of adult bilingual grammars, whether acquired early in childhood or later in adulthood, have led to observations that dual language exposure may result in local and selectively variable areas of grammar in one or both languages for an adult speaker (as compared to a native monolingual speaker). Much empirical work has focused on characterizing this divergence, both in developing and in apparently stable grammars, in order to a) determine if the underlying mechanisms for language acquisition are continuous from childhood to adulthood, and to b) identify the contextual factors that may lead to non-monolingual-like outcomes in mature grammars. This chapter presents empirical work examining the two groups of speakers to be examined in this project: first, (early) adult heritage bilinguals, and second, (late) adult bilingual speakers of a late-acquired L2. Specifically, I will review the following aspects of early and late bilingualism: 1) the nature of input and context of learning in bilingual acquisition contexts; 2) factors related to the development and maintenance of advanced linguistic outcomes in heritage and adult L2 speakers; 3) the acquisition of pronominal morphosyntax in contexts of Spanish-English bilingualism.

In discussing the causes and effects of bilingual language effects in early and late bilingualism, I will consider a variety of factors relevant to early and late language development: i) quantity, quality and mode of input; ii) the role of learnability in cross-language influence and transfer; and iii) the role of stabilization, attrition, fossilization and convergence in accounting for non-monolingual-like behaviour in bilingual grammars. A direct comparison between the two groups in question will be provided regarding these issues and the outcome of dual language exposure where empirical work has assessed both groups in a single study, placing special focus on the factors that comprise the macro-variable examined in the current study, age of exposure to bilingualism.
4.2 Age of Exposure in Early and Late Bilingualism

In comparing the two types of bilingualism examined in this study (2L1 and adult L2), the main variable being tested is age of exposure to bilingualism. There are two aspects to consider when examining this particular variable: 1) age as a biological factor of development (internal variable); and 2) age as a factor of learning context (external variable). In the following sections, I will review internal and external variables associated with age of exposure to bilingualism; this review will include the notion of a critical period in bilingual development, as well as the impact of naturalistic and classroom exposure on the process and outcome of bilingual attainment.

4.2.1 Age and a Critical Period for Bilingual Acquisition

Evidence of a biologically-determined critical period for L1 language development (Lenneberg, 1967; Newport, 1990) has had implications for how age is viewed in L2 development. A general observation has held that adults appear to have more difficulty than children in completely acquiring both phonological and morphosyntactic properties of a second language; consequently, an extension of a critical period for L1 acquisition to contexts of L2 acquisition has been proposed to account for this age-related difference (Bley-Vroman, 1989). According to this view, changes occur to the language faculty as a result of maturation; the individual reaches a critical moment in his or her cognitive development, and as such, his or her capacity to deploy the cognitive mechanisms used for the L1 to represent and acquire language is limited after this critical period has passed (i.e. puberty). In a generative-based study, such as the current one, this relates to the closing off of the availability of the innate language-specific cognitive mechanism thought to underlie L1 representation and development, represented most commonly by Universal Grammar (UG) and/or an innate language acquisition device (LAD). Two types of studies have supported such a maturational limitation: 1) those that have examined highly fluent L2 speakers first exposed to an L2 as adults and noted a general inability of this group to exhibit complete native-like knowledge of both phonological and morphosyntactic properties of the target language (Coppieters, 1987; Sorace, 1993; Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007); and 2) those that have illustrated a strong inverse relationship between adult native-like proficiency in phonological and morphosyntactic domains and age of first exposure to the L2 (Oyama, 1976; Johnson & Newport, 1989, 1991).
Of the latter group of studies, Johnson & Newport (1989; 1991) have most influentially argued for a biologically-delimited period of access to innate language-specific mechanisms for the acquisition of an L2. In examining a variety of morphosyntactic domains of the L2 English of native Chinese and Korean speakers, the authors showed that the adult bilinguals (L2 Eng-L1 Chin; L2 Eng-L1 Kor) who had been exposed to English as children (before 15 years of age) demonstrated consistent and constrained L2 knowledge as adults, with age of first exposure having an inverse relationship to native-like knowledge (up to 15 years of age). For those who had been exposed to English after the age of 15, this consistent and constrained L2 knowledge of adult bilinguals disappeared suddenly; this group of bilinguals showed highly variable, and somewhat random, knowledge of the same morphosyntactic domains. To the authors, this evidence suggested that a cognitive door had closed for those speakers who had been exposed to English after the age of 15; they had not and were not using the same cognitive mechanisms to acquire and represent their L2 as those who had been exposed to the L2 as children and young adolescents (i.e. before termination of puberty).

This evidence of a cognitive window of opportunity (Lee & Schachter, 1997) for the acquisition and representation of an L2 with an innate language mechanism (i.e. hypotheses\textsuperscript{35} of Bley-Vroman, 1989; Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007) has, nevertheless, faced criticism from three main sources. First, in a direct response to the evidence presented in Johnson & Newport (1989), Bialystok (1997) argued that the variable results of the bilinguals exposed to L2 English after the age of 15 was not as unconstrained as reported in the original study. In re-plotting the data of this group, Bialystok showed that the inverse relationship between age of exposure and proficiency that existed for the early sequential bilinguals continued as well for those exposed to the L2 after 15 years of age (the late sequential bilinguals); an age effect did exist, but there was no abrupt change in the constrained nature of the knowledge exhibited by L2 speakers, just a gradual decline in native-like proficiency with advancing age.

Second, this evidence of constrained but non-native-like performance has been demonstrated in a number of other studies examining speakers of a late-acquired L2. White

\textsuperscript{35} Fundamental Difference Hypothesis; Failed Functional Features Hypothesis; Interpretability Hypothesis, respectively.
(2003) provides a good review of such studies and concludes that, “There is now a considerable body of research whose results are consistent with the claim that learners arrive at mental representations for the L2 input which are systematic and UG-constrained” (p. 269).

Specifically, among this research are studies examining the same grammatical domain as Johnson & Newport (1991): late-acquired English L2 knowledge of wh-island constraints (also known as subjacency and/or locality constraints; Birdsong & Molis, 2001; White & Genesee, 1996). Birdsong & Molis (2001) examined the knowledge of wh-island constraints in the late-acquired L2 English of native Spanish speakers; if language-specific cognitive mechanisms were no longer available to adults to acquire an L2, then all adults acquiring the same L2 should be using the same general cognitive mechanisms to learn a new language, and as a result, have similar outcomes for the same grammatical domain. The results from Birdsong & Molis, however, pointed to a more native-like outcome for Spanish speakers of L2 English as compared to the Chinese speakers in Johnson & Newport. Given that, unlike Chinese, Spanish has similar wh-island constraints as English, the authors suggested that knowledge from the L1, developed and constrained by a language-specific mechanism (i.e. UG), was being used to analyze and constrain the representations of the L2.

Further, White (2003) also pointed out that the way in which the Chinese-English bilinguals examined in Johnson & Newport (1991) treated locality restrictions in English long-distance wh-sentences reflected grammatical knowledge typical of their first language, Chinese. In this way, the non-native-like knowledge exhibited by these bilingual speakers is not outside the scope of constraints permitted in natural language, they are simply typical of overall language knowledge. Persistent language transfer or cross-language influence at advanced levels of L2 proficiency has, thus, been one of the factors thought to contribute to divergent outcomes in late bilingual grammars, an outcome that should not occur if the innate language faculty responsible for development and representation of the L1 were closed off completely due to maturation after a critical period in or around puberty. If it were, developing and mature L2 grammars would likely look more random (individually, by language group and across language groups) than they do.

Finally, evidence against a strong critical period position for L2 acquisition has come from empirical work showing that while most late L2 speakers do not achieve native-like
monolingually-acquired targets for many phonological and morphosyntactic domains, there are some individuals who are able to achieve native-like knowledge for the typically challenging domains of a late-acquired L2. For example, while most English learners of late L2 Spanish exhibit persistent non-native-like divergence in their control of aspect and tense distinctions and gender agreement at advanced proficiency levels, Montrul & Slabakova (2003) and Montrul, Foote & Perpiñán (2008a) reported that of the most proficient late L2 Spanish speakers examined in their studies, some (20% and 3%, respectively) did fall within the range of language knowledge exhibited by control groups, that is to say native speakers of a monolingually-developed L1 Spanish. While non-native-like divergence may characterize the grammars of most speakers of a late-acquired L2 for selective grammatical domains, there is now sufficient evidence from individual results that counters the notion that adults are subject to strict language-specific biological limitations on the acquisition and representation of a late-acquired L2.

To summarize, the strong view of a biologically-determined critical period for L2 language development has been largely ruled out by i) studies showing that there is no sharp decline or abrupt change at any particular age or maturational point (such as puberty) in the language learning and representational capacity of an individual exposed to a late second language; ii) evidence of systematic linguistic constraints operating on late-acquired L2 grammars; and iii) individual results of late L2 bilinguals who have been able to completely acquire the typically challenging domains of the L2 grammar (with respect to the monolingually-developed L1 target). Nevertheless, evidence has supported a general tendency for age to be inversely related to an individual’s ability to reach monolingually-developed language-specific targets for a variety of phonological and morphosyntactic domains (see Montrul, 2008). This gradual age-related decline in language learning abilities has, therefore, led to weak maturational proposals for language acquisition across the lifespan: while language-specific cognitive mechanisms may, in principle, remain available through one’s life, other age-related cognitive changes may limit an adult’s ability to acquire and represent certain grammatical domains as a child would (Montrul, 2008). 36 Consequently, the earlier an individual is first (and consistently)
exposed to dual-language input, the more native-like (relative to a monolingually-developed target) he or she will eventually appear to be after a given period of time and a given context of exposure. While proposals of what exactly leads to this so-called sensitive period effect are varied, such a hypothesis has clear implications for the current study: those with an early exposure to two languages are expected to have a linguistic advantage over those who are exposed in adulthood to a second language, and as such, an early bilingual should demonstrate a more target-language-specific grammar (i.e. monolingually-developed) as compared to a late bilingual at a given level of proficiency in the same language.

4.2.2 Age, Exposure and the Context of Learning

While age may be a variable that has cognitive implications for the apparent declining language abilities in adolescence and adulthood, this variable also subsumes a variety of external factors that can have an effect upon both the process and outcome of language development in early and late bilingualism (Montrul, 2008, 2010). More specifically, if we compare the process and outcome of acquiring Spanish as a heritage L1 and an adult L2, we must compare the context in which the speaker is exposed to the language at the beginning and through the course of development in order to help us determine what the unique linguistic characteristics of fluent bilinguals tell us about the capacity to acquire languages at different ages.

Here, exposure to the target language will be identified as the input, defined as the raw linguistic data that the learner encounters through the course of acquisition (in speech and writing), and the resultant opportunities for communicative interaction that this exposure may provide. In the following section, I will consider the mode of exposure and the effects that both naturalistic and classroom exposure have been shown to have upon eventual linguistic outcomes for bilingual speakers. I will also consider these contexts of exposure from quantitative and qualitative perspectives; that is to say, the effect that the quantity and quality of exposure to developed) grammars (representation and use) for most domains of the target language. As such, the term sensitive period is preferred due to the strict limitations that the notion of a critical period represents.
input/language use has been shown to have upon eventual linguistic outcomes for the two types of bilinguals under investigation.

4.2.2.1 Context of Exposure in Early and Late Bilingualism

One of the most remarkable differences between adult heritage speakers and adult L2 speakers of a language is the context in which exposure begins and proceeds through the course of acquisition. In the case of the heritage speaker, exposure to the heritage language occurs as a typical L1, via naturalistic input/interactions in the home with parents and/or primary caregivers. Even if the child is exposed to an additional language in the very early stages of life (e.g. family members, childcare, etc.), the means by which a child has to develop a representation of the heritage language is through aural stimulus and verbal interactions. In this way, the language knowledge expressed by a heritage speaker (in development and as a mature speaker) is primarily related to the speaker’s early linguistic competence, or implicit linguistic knowledge (DeKeyser 2003).

On the other hand, adult L2 language learning in a foreign language setting (e.g., learning Spanish in Canada/US), the context upon which much empirical work in the field of adult L2 acquisition (at least in North America) is based, begins in a classroom. Right from the outset, adults begin to acquire language knowledge on the basis of three types of exposure: a spoken/aural stimulus, explicit grammatical instruction and a written/visual stimulus. These three types of exposure may vary substantially from class to class and program to program, depending on the chosen methodology for foreign language instruction; nevertheless, regardless of how much naturalistic exposure adult L2 learners have in the classroom, the immediate and perhaps lasting effect of writing and grammatical learning on adult L2 language knowledge cannot be ignored, especially when the grammatical domain to be acquired is optional and is explicitly taught (to some degree) in an instructional setting. In this way, the L2 knowledge exemplified by an adult (at any stage) may be the result of his or her implicit knowledge of the language (linguistic competence), but also related to a strong metalinguistic awareness they have of the rules (via explicit instruction) and structure (via visual cues) of the language itself.

In the following sections, the issue of language input/interaction will be considered for the two groups outlined above in terms of the mode, quantity and quality of language exposure that
occurs through the course of learning, and the effect that these three variables have been shown to have on both the process and outcome of early and late bilingual grammars.

4.2.2.2 Exposure in Heritage Bilingualism

A general observation of heritage language speakers is that they are a very heterogeneous group whose linguistic experience and ability vary greatly from speaker to speaker (Kondo-Brown, 2005; Montrul, 2010). Here, I will consider a group of heritage speakers who were exposed to both Spanish and English before the age of 2;0 and are communicatively fluent (advanced proficiency) in both languages as adults. In all cases, Spanish input/interaction occurred on the basis of communication in the home by native Spanish-speaking parents and through an extended family/community network, and English input/interaction occurred on the basis of communication with additional caregivers or family members (e.g., daycare, babysitters, siblings) and through interactions in the majority language of the surrounding community. Even though dominant exposure in the early developmental period for this group is expected to have been Spanish, these speakers would have had the experience of being simultaneous bilinguals, developing both Spanish and English as an L1.

As pointed out in a variety of studies in the area of early bilingualism (Paradis & Genesee, 1996; Genesee, 2010), quantity of exposure to input—as a measure of the number of waking hours exposed to and interacting in a language—is a variable that can affect the rate at which a young (pre-school) bilingual is expected to acquire particular grammatical domains, especially those that present some particular learnability problem for the child. For example, in the area of pronominal morphosyntax, bilingual children have been shown to exhibit persistent object omissions with obligatory transitive contexts for one or both of their languages as compared to monolingual children (Müller & Hulk, 2001; Belzil et al., 2007; Pérez-Leroux, Pirvulescu & Roberge, 2009). As noted in Chapter 3, one of the possible explanations that can account for this apparent delay to converge upon target-language-specific features at the same rate as monolinguals is that bilinguals have relatively reduced amounts of exposure in each of their languages as compared to monolinguals at the same stage of development; that is to say, compared to monolinguals, simultaneous bilinguals split their waking hours (to some proportion, not always 50/50) exposed to (and communicating in) different languages, and thus, may require
more relative time than monolinguals to resolve language-specific features of (apparently) variable/ambiguous systems.

Along with a relatively reduced quantity of exposure in each respective language in the pre-school years as compared to young monolinguals, young heritage speakers typically experience a pronounced shift in dominant patterns of exposure once they begin school around 4;0-5;0 years of age. Prior to school-aged years, Spanish heritage speakers typically experience dominant exposure to Spanish, allowing them to develop a good representation of early-acquired features of Spanish morphosyntax. Once these children begin school, however, this dominant exposure shifts to English through formal education and the on-going development of a social/peer network. Maintaining and further developing Spanish requires a consistent effort on the part of the family (and speaker) to insist on the use of the heritage language at home, among family members and in the extended Hispanic community (see Pérez-Leroux, Cuza & Thomas 2011b). Shifting dominant exposure from Spanish to English around 4;0-5;0 years of age has been shown to result in two types of language effects for certain grammatical domains (see Montrul, 2008, 2010):

1) language stabilization, whereby selective features completely acquired after 4;0-5;0 years of age (in monolingual development) remain incompletely acquired by the heritage speaker; and

2) language attrition, whereby selective features of completely acquired (early-acquired) grammatical domains are lost or altered in favour of a representation reflecting their shifting language exposure.

The degree of attrition that occurs through to adulthood are affected by a number of factors associated with patterns of exposure and use of the heritage language in childhood and adolescence (Montrul, 2008), factors examined previously in Chapter 3 and in more detail in section 4.3.

Finally, while it is often assumed that the quality of input (here, input as the template upon which speakers develop their own representations) to which a bilingual child is exposed represents standard use as exemplified by most native speakers, it has been recently highlighted that the language input to which Spanish heritage speakers in the U.S. are exposed (and presumably Canada) may in fact represent an already divergent form of the target language (Spanish), both qualitatively and quantitatively, and as such can result in “what seems to be
arrested development of the heritage language or attrition in adult bilingual knowledge” (Rothman, 2009, p. 156).

While few studies have directly tested this claim with evidence from child heritage speakers, the results from Paradis & Navarro (2003) tend to support the notion that non-target-like speech patterns in young child bilinguals may be the result of non-standard-like patterns of use exemplified by the native-speaking parents providing the input. In this case, while the increased overt use of subject pronouns in the Spanish of the young Spanish-English bilingual child was thought to have been due to a cross-language effect from English, Paradis & Navarro showed that the parents of the young bilingual used similar rates of overt subject pronouns in their speech, due to a dialectal constraint (Cuban Spanish). Consequently, both the quantity and quality of input may have an impact on the language knowledge exemplified by young bilingual speakers through the process and ultimately as an outcome of dual language development. Other studies examining the L1 attrition of native Spanish-speaking immigrants in the US also point to changing patterns of knowledge and use for certain grammatical domains in the L1 (Spanish), and as such, present evidence that the input to which first generation heritage speakers of Spanish are exposed may already be qualitatively different from the monolingual norm (Domínguez, 2009; Cuza and Frank, 2010).

4.2.2.3 Exposure in Adult L2 Acquisition

Like heritage speakers, fluent adult speakers of a late-acquired L2 are a linguistically heterogeneous group; however, unlike heritage speakers, who develop language knowledge almost solely on the basis of naturalistic and native (L1-developed) input/interactions, late L2 speakers of a language develop knowledge of the L2 under a wide variety of conditions.

In general, many L2 speakers of Spanish in North America are first exposed to Spanish in a language classroom. This particular context of learning provides initial language input via visual cues (written structure via texts and teaching grammatical rules), along with aural ones. In fact, in many academic-oriented language programs, the language in which the learner is provided instruction in the initial stages of learning (by the teacher and in the textbook) may be English, along with varying use of the target language. Further, students begin to develop production abilities only inasmuch as the program/teacher supports this type of classroom learning.
However, even when more relative class time is dedicated to communicative activities, late L2 learners typically communicate mainly with other learners of their communicative level, and as such, this context may not provide as good of an opportunity for learners to develop their oral skills.

Naturalistic input is available to language students in the classroom through a variety of means (teacher, audio exercises, videos, etc.) and tends to increase as the student advances levels; however, the type of aural input that Spanish students receive may consist of both L1- and L2-developed speech varieties, along with a wide variety of regional dialects. Especially in the area of phonology, but also in the area of morphosyntax, being exposed to a multitude of dialects through the developmental stage may provide ambiguous (and sometimes conflicting) evidence to the learner. For example, while speakers of most varieties of Spanish use an accusative pronominal clitic to refer to masculine and human direct objects, as in (70i), many speakers of a variety of dialects, including the Spanish of central and northern Spain as well as many varieties of Spanish-American dialects, employ the dative form of the clitic to refer to these male direct objects (and in some cases females too). This particular phenomenon is commonly known as leísmo and is illustrated in (70ii).

(70)  
  i. Lo pegué. (a mi hermano)  
      cl.acc. I hit (prt my brother)  
      “I hit him.”

  ii. Le pegué (a mi hermano)  
       cl.dat. I hit (prt my brother)  
       “I hit him.”

Students exposed to leísta varieties would, thus, not only be exposed to evidence of the standard variability of the Spanish pronominal system (related to case, gender, etc.), but would also be exposed to naturalistic input that adds variability (and ambiguity) to an already variable system (related to the semantic characteristics of the object). In the case of leísmo (and some other

37 This use of the dative form le or les to refer to direct objects is commonly referred to as leísmo aparente, ‘apparent leísmo’: the speaker maintains the distinction between accusative and dative objects, but uses the dative form of the pronominal clitic for the context described here. Leísmo real, ‘real leísmo’, is not as commonly attested where speakers have neutralized the pronominal system and do not make a distinction between accusative and dative objects (Fernández-Ordóñez 1999).
grammatical domains), the aural input heard may directly conflict with the rules taught explicitly in the classroom, a situation that may enhance an already challenging learnability problem in acquiring the pronominal system of Spanish.

Further to dialectal variation, L2 learners of Spanish are often subject to input provided by other non-native speakers of the target language (instructors, TAs, other learners). The type of naturalistic input to which learners are exposed, therefore, may already diverge from a language-specific use/distribution of certain grammatical domains (e.g. overt use of subject pronouns). It is not surprising, therefore, that late L2 learners do not converge upon monolingually-developed norms (control groups in empirical studies) for certain grammatical domains given the conditions that underlie much of a late L2 learner’s mode and quality of exposure to input/interactions.

In order to enhance their learning experience, many late L2 learners, thus, take advantage of opportunities to immerse themselves in a target-language environment, whether through stays abroad or through developing social networks in Hispanic communities. This experience may not only provide a more consistent quality of input, but also tends to increase the quantity of input and opportunities for communicative interaction (output). While it has been difficult to operationalize the notion of input quantity, the basic observation made for child bilinguals is relevant in the adult context: adults may need a sufficient quantity of exposure (input and output) in order to resolve a particular domain of grammar that manifests itself ambiguously or variably in the input to which they are exposed.

Recent results from a study on lexical gender have pointed to the role of exposure effects on the resolution of language variability for a lexical domain of grammar. Unsworth (2008) showed that the factor that most contributed to an English speaker’s accurate assignment of noun gender in L2 Dutch was length of exposure to input (Unsworth appears to mean length of exposure as her L2 participants were living in the target-language community). The speakers who had been living in the Netherlands for more than 8.5 years exhibited significant advantages in being able to accurately assign the more marked gender (neuter) to corresponding nouns. This variable had more of a significant impact on accurate gender assignment than both general language proficiency and age of first exposure to Dutch. Since those who had experienced the most input, or exposure, overall achieved the most native-like grammars—regardless of their age
of exposure or relative language proficiency—then quantity of exposure to input, or rather, the relative time exposed to and communicating in the language, could be considered an important factor in the resolution of lexically-determined variability.

4.2.2.4 Naturalistic vs. Classroom Learning

As mentioned, one of the main factors that differentiates the language learning experience for early heritage bilinguals and late adult L2 bilinguals in North America is the context of exposure to the language: naturalistic exposure for heritage bilinguals and a mixed exposure with early predominance of classroom exposure for adult L2 bilinguals. While few studies examining acquisition within a generative framework consider the context of learning on the outcomes of early or late bilingual development, the distinction between these two types of learning does appear to have an impact upon the way that these two groups of bilingual speakers attend to tasks that are designed to empirically test the way they represent and use language. Montrul, Foote and Perpiñán (2008a) tested early and late bilingual speakers of Spanish for gender assignment and agreement on three separate tasks: 38 1) a Written Picture Identification Task; 2) a Written Gender Recognition task; and 3) an Oral Picture Description Task. Results showed that for both written tasks, late bilinguals outperformed the early bilinguals, 39 while for the Oral Picture Description Task, the early bilinguals outperformed the late bilinguals. 40

While in other L2 studies, differential results for late L2 bilinguals on receptive and production tasks have often been accounted for through a competence-performance asymmetry (i.e. Haznedar & Schwartz, 1997; Lardiere, 2000; Prevost & White, 2000), Montrul, Foote & Perpiñán suggest that the late bilinguals’ relatively greater success on the written tasks than the

38 The authors note that error rates above 10% are interpreted as some deficit or failure in controlling either gender assignment (a lexical domain) or agreement (a syntactic domain) for bilinguals based on previous findings that adult monolinguals rarely make any gender errors (assignment or agreement) (535-536).

39 Late bilinguals vs. early bilinguals for mean proportion of accurate responses on each written test, respectively: 90% vs. 85% and 89% vs. 83%. Both experimental groups had similar ranges of accuracy for each written test: 60-100% for the Written Picture Task and 37-100% for the Written Recognition Task (compared to 97% and 99% for native monolingual speakers, respectively).

40 Early bilinguals scored a mean proportion of 89% of accurate responses (58-100% range) against the late bilinguals’ 72% (50-100% range); native monolingual controls scored a mean accuracy rating of 99% (98-100% range) .
oral one, and their relatively higher rates of accuracy on the written tasks as compared to the early bilinguals, may be the result of having greater metalinguistic awareness, mainly due to the effect of explicit learning that occurs in classroom contexts. Conversely, the early bilinguals’ generally consistent results on all three tasks (83–89% mean accuracy), and their relatively greater accuracy than the late bilinguals on the Oral Production task (89% vs. 71% accuracy, respectively), points to this group having greater linguistic control of the lexical and syntactic mechanisms that underlie gender assignment and agreement in Spanish. These results, while reinforcing the relationship between implicit learning and naturalistic exposure and explicit learning and classroom exposure, do not imply that there is no implicit learning in classroom contexts, nor that heritage bilinguals have no metalinguistic awareness. What these results point to is simply that those with more formal and explicit exposure to language in a classroom, like late L2 speakers of Spanish, may resort to explicit knowledge of their L2 for more formal tasks when their implicit knowledge is unreliable, as compared to those who have not had as extensive exposure to formal language learning, like early heritage speakers of Spanish. Given these results, therefore, it may be difficult to determine whether a competence-performance asymmetry really exists for any type of bilingual speaker if receptive and production knowledge are tested in different modalities (visual/written vs. aural/oral), or if the representation of a grammatical domain varies according to the mode of production (see Rothman, 2007 for discussion on inflected infinitives in Brazilian Portuguese).

While the supposed relationships between implicit learning and naturalistic exposure and explicit learning and classroom exposure may have task-related implications (see also Gass, Mackey & Ross-Feldman, 2005; Rosa & O’Neill, 1999), the question remains as to whether or not one of these types of learning, and/or contexts of exposure, confers an advantage to the language competence of bilingual speakers in the long-term. Results from studies that have examined the issue of naturalistic vs. classroom learning have generally highlighted the benefits of classroom learning for late beginner-intermediate adult L2 learners. The type of form-focused instruction provided in many language classrooms (i.e. explicit instruction, corrective feedback, etc.) may help these speakers at earlier stages of acquisition generalize rules and test language-specific hypotheses of their L2 for selective grammatical domains at a faster rate than those who have not had exposure to form-focused instruction (Mackey, 1999; Gass, Svetics & Lemelin, 2003; Pavesi, 1986). Nevertheless, if viewed from an input-output perspective, these perceived
benefits of classroom instruction may be limited in contributing towards native-like outcomes in adult L2.

First, while the results from Pavesi (1986) showed that Italian learners of L2 English with predominantly formal instruction used more advanced, marked forms of relative clause types than Italian learners of L2 English without formal instruction, this apparent classroom learning advantage may be related to the diversity of input available to each group of speakers. More marked types of relative clauses (i.e. object of preposition, genitives) tend to be used less often in speech than the unmarked types (i.e. subject and object relatives), and as such, presenting the more marked structures in a classroom may enhance the relative frequency (quantity) that these appear in overall input. The output, therefore, of both groups of learners may be more reflective of the relative proportions of experience with each type of relative clause at a given moment of learning/exposure, rather than classroom learning (explicit and form-focused instruction) having a specific acquisition advantage over naturalistic learning in the long-term.

Second, overall results from comparing classroom to naturalistic learning have demonstrated that explicit teaching alone does not enhance L2 learning, but rather enhancement is seen when explicit learning is coupled with target-language input and an opportunity to practice the form (Ellis & Laporte, 1997). Again, quantity of exposure for particular domains may be enhanced in the classroom in a way that wouldn’t normally be in naturalistic environments at a given moment of learning. If a speaker is exposed to only naturalistic input consistently and robustly, the relative proportion of exposure to any one grammatical domain may be more or less consistent throughout their learning process, whereas in a classroom context this proportion of exposure may be exaggerated or enhanced for a certain period where specific pedagogic focus is placed on a particular grammatical function, a condition that may lead to enhanced use of this feature in the output. Further, some evidence has also pointed to the features of the target language to which students are exposed in a formal classroom setting (as those with a typical L1 experience and L2 classroom experience will have) is different than that which those are only exposed to the vernacular will experience (as those with a typical heritage language experience) (Rothman, 2007). In the long-term, therefore, these two groups may appear to have different ways of representing and using a grammatical domain due to the relative presence or absence of certain features of that domain due to the formal or informal nature of exposure.
Classroom and naturalistic exposure, therefore, may lead to (perceived) differential outcomes in bilingualism because learners may be exposed to different quantities and qualities of exposure for any particular grammatical domain at any particular moment in the language-learning process. In that regard, exposure to standardized rules or a neutralized dialectal form in classroom learning can also lead to stabilization effects in the long-term if the input or rules explicitly taught in the classroom underestimate what must be acquired, or the input and rules to which the speaker is exposed contradict each other.

4.3 Development and Attainment in Early and Late Bilingualism

Recall from Chapter 3 that I adopt a learnability view of bilingual language behaviour; that is to say, the bilingual effects seen in developing and mature bilingual grammars are most likely to be found for domains of grammar that are underrepresented in input and/or that are expressed variably/ambiguously from a morphosyntactic point of view. Unlike monolinguals, who typically resolve learnability problems over time and converge to the most target-specific grammar, bilinguals have more diverse linguistic resources at their disposal to resolve a learnability issue or express themselves in the most clear morphosyntactic way. Bilingual-like behaviour is, therefore, local and selective on the basis of a variety of factors, and may persist according to these same factors, such as how congruent the languages are from a morphosyntactic perspective (underlying grammar), the mode of exposure, the quality of input, and the (shifting) relative quantity of input/interaction in the languages to which the speaker is exposed at any point in his or her life (dominance issues). In the following sections, I will review the studies that have aimed to account for the process and outcome of early and late dual language exposure, and review the empirical data regarding the knowledge and use that early (heritage) and late (adult L2) bilinguals have of the more challenging domains of grammar. To conclude this chapter, I will provide a review of studies that have specifically looked at the acquisition of Spanish pronominal morphosyntax in contexts of Spanish-English bilingualism (heritage and adult L2 speakers) in order to make predictions for the current study.
4.3.1 Simultaneous and Heritage Bilingualism

4.3.1.1 Early Childhood

Some of the first formal proposals regarding the development of language in young bilinguals used evidence of language mixing in speech production to support the notion that bilingual children begin acquiring more than one language with a common linguistic system. Volterra and Taeschner (1978), for example, proposed the Unitary Language System Hypothesis (ULH), which stated that bilingual children start off with a common language system and gradually differentiate the system around age 3;0, when their experience with language input is robust enough to warrant the instantiation of differentiated grammars. Once the production of young bilinguals was analyzed more closely, however, the ULH was discounted in favour of a theory proposing that bilingual children construct differentiated language systems from the earliest stages of dual (or multiple) language exposure (Genesee, 1989; Meisel, 1989).

The Dual Language System Hypothesis (DLSH, Genesee, 1989) was initially proposed and has been consistently supported on the basis of evidence that bilingual children (younger than 3;0) exhibit the same timelines of development for certain grammatical properties in distinct languages as developmentally-matched monolingual children of the same language. For example, Meisel (1989) examined the word order and subject-verb agreement patterns of two French-German bilingual children between the ages of 1;0 and 4;0. If bilingual children were using common syntactic rules for both languages in the early developmental period (2;0-3;0), they would be expected to show evidence of a common word order, either SVO (like French) or SOV (like German), for utterances in both languages, unlike the respective monolingual children during this period. Contrary to what the ULH would predict, Meisel reported that these bilingual children maintained language-specific word order patterns in each of their languages from 2;0 to 3;0 (SVO for French and SOV for German) that conformed to the developing language exhibited by monolingual children for French and German, results that supported the DLSH. Further

---

41 Formalized as the Unitary Language Hypothesis (ULH), this theory was constructed on the basis of three stages for the lexical and grammatical development in early bilingualism. Initially in Stage I (<~2;0), bilingual children instantiate both a common lexicon and syntactic system for both languages; in Stage II (~2;0-3;0), their lexicons become differentiated, but they continue to use a pooled resource of syntactic rules; and then by Stage III (around age 3;0), bilingual children have enough experience with each language to construct distinct lexicons and grammatical systems for each language to which they’re exposed.
studies examining the early bilingual development of other grammatical properties such as negation, finiteness, noun phrase syntax and subject syntax for a variety of language pairings such as Dutch-English, French-English, Spanish-Basque, English-Catalan and English-Inuktitut (De Houwer, 1990; Paradis & Genesee, 1996, 1997; Barreña, 1997; Juan-Garau & Pérez-Vidal, 2000; Zwanziger, Allen & Genesee, 2005) have also supported the notion that bilingual children have differentiated language systems from the earliest stages of development.

Nevertheless, despite the evidence pointing to language differentiation in early bilingualism, it has always been clear to most observing bilingual children that they are not simply “two monolinguists in one” (Grosjean, 1989). Even though the ULH was ruled out by evidence provided by the studies mentioned above, a strong version of the DLSH could not account for evidence that some bilingual children do things with language that their monolingual counterparts do not; that is to say, in some cases bilingual children exhibit unique patterns of language use for one or both languages not attested in monolinguals at the same stages of development. The aim of research examining bilingual language effects in early bilingualism has, thus, focused on determining the types of effects exhibited by bilingual children (i.e. domains of grammar and constraints) and the factors that lead to these bilingual effects in early bilingual grammars.

In general, bilingual language effects in childhood are local and selective; that is, a) they do not occur generally across all grammatical domains for a language pairing; and b) they occur for a particular grammatical domain in some language pairings, but not others. Typological relatedness has, thus, been pointed to as an important factor in predicting where divergence may occur in early bilingual grammars (Tracy, 1995). For example, while a number of studies have shown that bilingual language effects may occur in child grammars for the domain of subject syntax between a null-subject and a non-null-subject language (Serratrice et al., 2004; Haznedar, 2007), this does not appear to occur for all language pairings. Zwanziger, Allen & Genesee (2005) followed six English-Inuktitut bilingual children ranging in age from 1;8-3;9 for approximately one year and showed that all bilinguals fell within target-like ranges for language-specific knowledge of subject omissions in each of their languages, despite the fact that Inuktitut is a null-subject language and English is not. Although no concrete explanation was made regarding the underlying reasons for this differentiation, the authors mentioned that the
morphosyntactic or typological divergence between the two languages (English as an Indo-European language and Inuktitut as an Eskimo-Aleut language) was a possible explanation for the lack of universal bilingual language effect in the domain of subject syntax.

Further to the observation that bilingual divergence is local and selective, investigators have pointed to two basic types of bilingual language effects in early childhood grammars: qualitative and quantitative effects. In the first case, bilingual children appear to use structures in one of their languages that monolingual children of the same language do not use, resulting in what often appears to be a direct influence, or transfer, from one language to the other. Specifically, as noted in Chapter 3, this type of bilingual language effect has been observed in the developing language of young bilinguals whose respective languages exhibit asynchronous timetables of development for the particular grammatical domain involved (Gawlitzek-Maiwald & Tracy, 1996; Döpke, 2000; Yip & Matthews, 2000). In all cases, the bilingual children appeared to transfer representations from one language to express in the other language a representation he or she had yet to acquire. Nevertheless, this type of interaction did not appear to persist past the stage at which the child was expected to acquire the particular structure in the slower developing language; that is, once the child reached the stage at which monolingual children were expected to acquire the particular domain, they no longer exhibited the qualitative effect. This type of bilingual effect is, thus, not expected to persist into late childhood or adult bilingual grammars as long as input levels for both languages remain relatively robust past the stage of early grammatical development (i.e. approximately 3;0+).

Unlike the cross-language qualitative effects seen in the studies described above, other empirical work on young early bilinguals have shown that these bilinguals tend to exhibit cross-language interactions that result in quantitative effects for some grammatical domains. During the course of L1 development, young bilinguals have been shown to exhibit some sort of delay in reaching target-like representations for certain grammatical domains on the same timetable as their monolingual peers. One explanation of this apparent delay is that these bilinguals have been exposed to a relatively less quantity of input in each language by a particular age/moment of development as their monolingual peers, and as such, may maintain early developmental representations (i.e. UG default representations) longer than monolinguals (recall the Default Representation Hypothesis, Pérez-Leroux, Pirvulescu & Roberge 2009). Hulk & Müller (2000)
examined the rate of object drop in the speech of a Dutch-French bilingual child and a German-Italian bilingual child. They found bilingual effects, with bilingual children exhibiting higher rates of illicit object drop in the more variable languages, French and Italian, at the same stage of development as their monolingual peers. No empirical results were provided for the German and Dutch of these bilinguals, who were assumed to have instantiated target-like rates of obligatory object use in the Germanic languages. Belzil et al. (2007) found similar results for object syntax with French-English bilinguals, but found higher rates of illicit null objects in both the languages of the bilinguals tested; the authors accounted for the results by proposing that bilinguals need longer time to resolve ambiguities related to object use because of their relatively reduced exposure to input in each respective language (as compared to monolinguals), and as such, maintain a developmental (UG) representation for objects (the null referential representation) longer than monolinguals in both languages.

The instantiation of a developmental representation in (one or both) bilingual grammars past the stage at which it is typically resolved in monolingual children is considered here as a general bilingual effect, and is not expected to be evident in the language of mature early bilingual speakers. Nevertheless, in cases where input in one language is not robust enough to resolve a target-language-specific representation for a late-acquired domain, or the non-target language provides a strong representation for that domain, then this type of developmental delay may lead to bilingual interaction effects. In general, there are three ways that young bilingual children may resolve a (persistent) developmental grammar:

1) Through convergence to the target-language-specific grammar (through sufficient (mode/quantity/quality) exposure to the target language);

2) Through a stabilization of the developmental representation, a representation that crucially must be interpretable in the target grammar (for the grammar and for the interlocutors);

3) Through the overall language knowledge that the child has (cross-language resources).

In this way, the developmental representations seen in child (2)L1 grammars may persist past early childhood, but dual-language exposure and/or the availability of cross-language resources may have the effect of altering the way that a bilingual child interprets/uses a particular grammatical domain at any one moment, or even in the long-term.
4.3.1.2 Mid-Late Childhood and Adolescence

For a child growing up in an English-language community, who has spoken another language from birth (L1 in contexts of immigration, home language (L1; 2L1) in contexts of heritage bilingualism), there may be a variety of linguistic consequences that accompany the shifting patterns of exposure related to the beginning of full-time school around 4;0-5;0. In mid-late childhood and adolescence, therefore, bilinguals who have been exposed to a home (heritage) language an English before the age of 4;0, or who immigrate around mid-childhood, may begin to shift from strong linguistic abilities in their L1/home language (vernacular, spoken language) to strong linguistic abilities in English (vernacular, spoken language, literacy, formal/academic abilities). Those who have studied this population more extensively have proposed that both stabilization and attrition effects may occur through the bilingual’s childhood and adolescence to varying degrees for their L1/heritage language, which, thus, results in a population of early adult bilinguals with a wide range of proficiency levels and linguistic abilities in the minority language.

Polinsky (2008a; 2011) examined the language abilities of Russian child and adult monolinguals and child and adult heritage speakers of Russian in the US on their narrative abilities and interpretation of relative clauses. She found that while monolingual Russian children performed like monolingual adults on all tasks, child heritage speakers of Russian converged more to the monolingual baseline provided by these monolingual control groups than adult heritage speakers. She interpreted these results to suggest that while child heritage speakers (around 6;0) have had robust enough input to converge to monolingual-like norms, they generally experience attrition for a number of morphosyntactic domains (case distinctions, gender agreement, pronominalization rules, reflexivization rules) and often arrive at adulthood with reanalyzed grammars (i.e. more simplified than that of Russian monolinguals and/or converged bilingual grammars). In Polinsky (2008b), the author further tested to see if relative proficiency in adulthood affected the degree of attrition for Russian case and gender systems. She found that adult heritage speakers of Russian with more advanced proficiency exhibited a three-way case system (monolingual-like), while those with lower proficiency levels had a two-way case system. Further, while she found that all heritage speakers had a good representation of gender, she found that those with lower proficiency produced more agreement errors in production that those with
higher proficiency levels. Overall, Polinsky’s (and others) work on Russian heritage speakers in
the US suggests that heritage speakers of a language will experience different degrees of
language stabilization and loss as they pass through the late childhood and adolescent years,
processes that will lead to apparent bilingual effects in adult grammars, and as such, apparently
different degrees of proficiency in the heritage language. Here, proficiency was measured in
terms of lexical size and speech rate; as such, as suggested by Schwieter & Sunderman (2008), a
bilingual speaker must have both good lexical knowledge and a fast and efficient lexical access
and selection process in order to maintain high levels of proficiency through to adulthood.
Consequently, it appears that those heritage speakers who both have consistent/expanded
exposure to the heritage language in childhood and adolescence, as well as active verbal
interaction in the heritage language over these formative years will arrive at adulthood with
relative greater proficiency (both passive and active), and consequently less bilingual effects for
the most vulnerable domains of grammar.

Further to the research by Polinsky, Jia and colleagues (Jia, 2004; Jia & Aaronson, 2003;
Jia et al., 2006; Jia & Fuse, 2007) have conducted research on the L1 maintenance/loss (and L2
acquisition) of child immigrants to North America. While age of arrival (AoA) is generally a
good predictor of L2 attainment and L1 loss (earlier arrival = more L1 loss/better L2 attainment),
the authors have shown that younger arrivals are not necessarily guaranteed complete attainment
in the L2 nor more loss of the L1, nor will older arrivals fail to converge to L2 targets and
maintain all aspects of their L1. Jia (2004) notes, “cognitive, social and cultural factors jointly
shape the course of L2 acquisition and L1 maintenance by immigrant children and adolescents”
(p. 368). These studies have collectively (and consistently) shown that environmental factors
such as language/culture of origin (i.e. European vs. Asian), motivation to learn and attitudes
about the host language, language preferences, and how rich their L2 environment is affect the
relative level of maintenance of the L1 to adulthood. However, if one considers notions like
motivation, attitude and language preference, these are all factors that may relatively increase or
decrease the speaker’s exposure and use of that language on a regular basis. Consequently, while
these non-linguistic variables are a good way to “gain traction” (Bohman et al., 2010) in language
development or maintenance through to adulthood, these factors (having good motivation to learn
or a good attitude to the language in question) are not necessarily directly related to linguistic
abilities or overall proficiency; the most deterministic factor that appears to be related to more
complete outcomes in adulthood for the minority language of a young bilingual is the degree of exposure and interaction/use in the minority language.

Pérez-Leroux, Cuza & Thomas (2011b) examined the linguistic and attitudinal factors associated with contexts of Spanish-English bilingualism in Toronto, Canada. This study tends to support the notion that the transmission of a minority/heritage language in an English-majority community, and consequently the linguistic abilities of young minority/heritage speakers (4;0-8;0), is mostly determined by the degree of exposure and interaction/use of the target language. In this study, while all families had positive attitudes to Spanish and the notion of bilingualism, and all child participants exhibited a positive view of Spanish and their language heritage, those families with the most positive attitude towards the maintenance of Spanish had the children with the lowest proficiency levels in Spanish overall (narratives and linguistic measure of pronominal morphosyntax). The parents of these children tended to be highly educated, middle-class professionals, highly proficient in English, who were proud of their first language and culture, valued their heritage and had a strong desire that their children speak Spanish fluently as adults, but lived in more middle-class Anglo communities, and typically didn’t insist on a strict Spanish-only rule in the household or in interactions with their children. Those children who were more highly proficient in Spanish (and even exhibited dominance in Spanish as children 4;0-8;0) came from families who were more recent immigrants living in a more dense community of other immigrants (including Spanish-speaking immigrants), and who had more neutral attitudes towards Spanish and bilingualism. The parents of this group were variably educated, with varying abilities in English, but in general, these parents were less likely to communicate with their child if he or she spoke in English (for functional or social reasons). Consequently, these children appeared to have greater relative exposure to Spanish, and importantly, used Spanish more in their own regular communication (with parents, extended family, siblings and neighbours) than the previous group of children. Overall, therefore, regular exposure and use appear to be strong deterministic factors in the transmission and maintenance of a minority/heritage language in mid-childhood, and most certainly through adolescence.

4.3.1.3 Linguistic Outcomes of Heritage Bilingualism

As noted in section 4.3.1, adult early bilinguals who live in an English-dominant community are a very heterogeneous group in terms of linguistic ability and overall proficiency in the
minority/heritage language (Montrul, 2008; Kondo-Brown, 2005). Much of the experimental work on adult speakers of a heritage language, therefore, can generally be divided into two groups: those looking at speakers of low-intermediate proficiency levels, and those looking at speakers of advanced-near-native proficiency levels. Further, there appear to be two main types of empirical studies that include heritage speakers as experimental groups:

1) those looking to answer questions regarding the notion of vulnerability in bilingual grammars and the factors (linguistic, psychological, sociolinguistic) that lead to apparently incomplete grammars in adulthood; and

2) those looking to answer questions regarding the notion of a critical period in L2 acquisition, and as such, compare the knowledge and use of certain (problematic) grammatical domains by both L2 speakers and heritage speakers of a language.

Most of the empirical work from North America examining the outcome of a heritage language experience (i.e. adult speakers) has mainly (but not exclusively) comes from the study of three languages: Russian, Spanish and Korean. Given that these languages are very different typologically and vary in their typological distinctions with English, we now have a good body of research to be able to draw some generalizations about what a context of heritage acquisition means in the long-term for these speakers, what we can predict linguistically for them, and, thus, what we can expect from them and future generations of bilingual speakers in North America.

In general, most research has shown that heritage speakers exhibit non-monolingual-like knowledge and/or use of morphosyntactic domains that are traditionally problematic from a learnability perspective. In the case of Spanish, Montrul and colleagues (Montrul, 2002; 2004a; 2004b; 2006; 2009; 2010; Montrul et al., 2006; Montrul & Bowles, 2009; Montrul, Foote & Perpiñán, 2008a; 2008b; Montrul & Ionin, 2010) have shown that heritage speakers exhibit non-monolingual-like behaviour in production and/or receptive tasks testing aspect (preterite and imperfect), mood (indicative and subjunctive), underrepresented/variable aspects of pronominal

42 Here, I will focus on the experimental work that has tested early bilingual populations in North America. Note that much of the foundational work in the field was established on the basis of observations of speech patterns in oral production of adolescent and adult heritage speakers. The relatively new field of study focussing specifically on an experimental analysis of heritage/minority language speakers is due in large part to the observations and analyses made in these early studies, for example Silva-Corvalán (1994; 2003). Further, empirical research is also ongoing in Europe, with a good representation of work on Italian and English as a minority language (see work and reviews in Sorace & Serratrice, 2009 and Sorace et al., 2009).
morphosyntax (i.e. clitic-left dislocation, also see Zapata, Sánchez & Toribio, 2005), direct object marking (DOM), the morphosyntactic properties of unaccusative-unergative verbs, gender agreement and the interpretation of definite determiners in generic contexts. Relatively higher proficiency appears to be related to more complete outcomes for most of these domains (compared to a monolingual baseline). This is true for other languages as well, like Russian and Korean, where there seems to be a graded effect in terms of how target-language-specific heritage speakers perform experimentally according to relative proficiency level and as compared to monolinguals and/or L2 speakers (Polinsky, 2006; Suh, 2007). Further, when comparing heritage speakers to late L2 speakers, the heritage speakers sometimes (not always) exhibit an advantage over L2 speakers, but don’t necessarily converge to monolingual-like targets for those domains (Montrul, 2002, 2004a; Bruhn de Garavito, 2002).

Despite heritage speakers exhibiting an advantage for some domains, more recent work has suggested that this advantage is often limited to particular language modes. While heritage speakers often exhibit advantages on production tasks and those that tap linguistic intuitions on the relative grammaticality/ungrammaticality of certain domains in mostly oral/aural tasks, L2 speakers appear to have an advantage on tasks where an academic or metalinguistic awareness aids in the assessment of target language structures (Montrul, Foote & Perpiñán, 2008a; Rothman, 2007; Kim, Montrul & Yoon, 2009). Consequently, heritage speakers often perform more like monolingual speakers on production/intuitive activities where explicit grammatical knowledge or a more formal/academic register does not facilitate the completion of experimental tasks, while L2 speakers often perform more like monolingual (educated) speakers on written/paper-and-pencil tasks where they are being tested on grammatical domains for which they have explicit grammatical knowledge and/or formal academic language training. Again, more complete outcomes (i.e. relative to monolingual baselines) are typically related to relative levels of proficiency: the higher the proficiency the more monolingual-like the bilinguals’ performance is.

Overall, task-related effects showing an advantage in production to heritage speakers—or rather better intuitive knowledge, and an advantage in formal tasks to L2 speakers—in their use of more educated/explicit grammatical knowledge, change according to the proficiency levels of the participants: those with higher proficiency (however measured) typically exhibit more
monolingual-like behaviour. Here, bilingual speakers typically exhibit non-monolingual-like effects for similar grammatical domains (those where there is some ambiguity/variability/underrepresentation of the particular feature/domain in input), but the speakers also show different degrees of incompleteness depending on the interaction among age of exposure, the grammatical domain (cross-language typologies) and the mode of performance.

4.3.2 Adult L2 Acquisition

4.3.2.1 The Initial State and Process of Adult L2 Development

White (2003) highlighted the importance of considering the so-called initial state in the examination of adult L2 acquisition: first, knowing how adults start the acquisition process will help in understanding what the course of development will look like; and second, characterizing the initial state can provide a framework for how successful adults can be in the acquisition of a second language (pp. 58-59). While bilingual children may be perceived as beginning the dual language learning process as blank slates—or rather, according to a generative view with a full complement of universal grammatical features and operations available to build independent grammars—adult monolinguals have already constrained a set of features and operations to instantiate a language-specific grammar for their L1. While the aim here is not to review all proposals that deal with the initial state and developmental process of a late-acquired grammar (see White, 2003, Gass & Selinker, 2008 for reviews), I will comment on the nativist positions that have some consequences for the way that cross-language effects may occur at early stages of development and proceed as the speaker becomes more balanced as a bilingual over time.

I adopt a strong nativist position here of L2 acquisition where an adult has access to all of the linguistic information and mechanisms for analysis that are available to the child speaker; the difference here is that, unlike young children, adults have already established a well-formed lexicon and a certain degree of automatization in the system related to the ways that they formulate linguistic expressions for a particular semantic expression. As such, as mentioned previously, the adult who begins to acquire a new language does so as a very unbalanced bilingual: his or her L1 representations and language processes are much more robust than those of the L2. Empirical work in the field of L2 acquisition has generally established, however, that adults can overcome/by-pass the language knowledge and processes of their L1 to acquire
features and properties of grammar of the L2 that are not instantiated in their L1. In the case of English and Spanish, English speakers have been shown to be able to acquire the null-subject status of Spanish (Al-Kasey & Pérez-Leroux, 1998), the gender system of Spanish (White et al., 2004), and the pronominal clitic system of Spanish (Liceras, 1985) even though the features associated with these grammatical domains are not native to the L1.

Nevertheless, it is obvious in listening to a new speaker of Spanish that he or she does not speak like a child (that UG is the starting point for development), but rather speaks more like an English speaker speaking Spanish (grammatically-speaking). While it may be obvious that the English speaker begins speaking Spanish by transferring his or her English to Spanish, there has been some disagreement over what exactly is transferring from the L1 in cases of late L2 acquisition. There are three basic approaches that attempt to account for the empirical evidence of early L2 knowledge. First with the Minimal Trees hypothesis, Vainikka & Young-Scholten (1994) suggested that the substantive categories (N, V, Adj, etc.) transfer from the L1, but that the learner has to acquire all functional grammar directly on the basis of exposure to L2 input (and in theory, in searching the UG). Here we have a theory that attempts to account for the reason that functional categories appear to be the most challenging to acquire for L2 speakers (and all speakers in general); however, there is quite of bit of evidence from early L2 stages that speakers may also transfer certain functional categories from their L1 in formulating an L2 linguistic expression (White, 1990). The Valueless Features proposal of Eubank (1994), thus, helps to resolve this issue by suggesting that both functional and lexical categories transfer from the L1, but that their settings are inert; that is to say, that the speaker needs input from the L2 to set the language-specific features associated with those categories. This proposal attempts to account for evidence of the apparent randomness or optionality that L2 speakers exhibit in the early developmental stages when attempting to apply verb inflections and other functional grammar. The trouble with this proposal, however, is that not all speakers of all language pairings exhibit a random grammar at the early stages of development for the same grammatical domains (Yuan, 2001, cf. White 2003).

The Full Transfer/Full Access (FTFA) proposal of Schwartz and Sprouse (1996), therefore, attempts to account for all L1-related evidence of transfer, but also all evidence that L2 speakers can, in principle, acquire a complete grammar for any late-acquired language (on the
basis of their universal language knowledge). While this theory covers more or less much of what could ever happen from an empirical point of view, it also works well from a systematic point of view if we consider the language model outlined for bilingualism in Chapter 3. If L2 speakers begin their acquisition process by building morphonological information related to the new language to which they’re exposed, their first grammatical expressions (2 words+) may be formulated on the basis of their L1 morphosyntactic information (the lemmas). As such, at the beginning stages of acquisition, many L2 speakers appear to be lexicalizing an L1 frame with L2 words. Once they are more exposed to the L2, these speakers may begin to acquire the target-language-specific morphosyntactic features associated with the L2 lexical entries that have begun to emerge in the lexicon of the speaker. As the FTFA proposes, where the L1 and L2 are congruent, the speaker really doesn’t have to learn anything; the morphosyntactic analysis of the L1 satisfies the representational needs of the L2. However, where there are representational differences, the speaker must be able to both perceive the difference and then go searching for a way to resolve the difference. Here, UG may provide the speaker with the necessary information to resolve the representational issue, or may provide intermediate representations (an interlanguage grammar, Selinker, 1972) until the target-language representations are acquired. So, in terms of the representational issue, late L2 learners should, in principle, be able to acquire a complete array of target-language-specific morphosyntactic features (lemmas), as long as they are able to figure out what they have to learn from their input. Here, we can see how an optional or more variable L2 grammar may affect an L2 speaker’s ability to converge on the target-language-specific representations, especially if the representations in their L1 are invariable and somewhat overlap (at least are syntactically congruent) with the L2.

Further to the representational issue, under the language model outlined in Chapter 3, L2 speakers must begin to be able to recruit the function of executive control to begin to effectively inhibit the grammar of their L1 when speaking in the L2. Having a representational or a processing gap may require these speakers to use their overall language resources (including their L1) to formulate a linguistic expression in the most monolingual-like way possible in their L2. As such, the function of executive control may work variably according to a functional need the speaker has to use his or her overall language knowledge to formulate a linguistic expression. There is, therefore, another issue to consider in the early stages of L2 acquisition in addition to the consideration of which linguistic properties/features L2 speakers have available and use: the
degree to which they can recruit the attentional control for linguistic purposes. Here, more traditional studies in the field of L2 acquisition (generative or applied fields) do not directly address this issue. In applied research, there is some information regarding how automatized an L2 grammar is (or isn’t), but most times this discussion doesn’t go beyond the consideration of a general description of procedural and declarative processes and the role of output in L2 development (Swain, 2005). Bilingual work in the field of psycholinguistics, therefore, can help us better understand what L2 speakers need to do in order to develop a system that allows them to communicate in the most monolingual-like way. As mentioned previously, some work has been conducted on bilingual populations of different proficiencies, where proficiency is defined according to lexical size (vocabulary and functional grammar) and efficiency (speed) of lexical access (Costa & Santesteban, 2004; Costa et al., 2006; Schweiter & Sunderman, 2008). In general, these studies have shown that inhibitory control is important in low to advanced levels of language proficiency in late bilinguals in order to effectively inhibit the strong activation of their relatively stronger grammar (their L1). However, very advanced L2 speakers appear to be able use target-language-specific resources more directly without resorting to the use of as much attentional control. Consequently, these very high proficiency L2 speakers have enough exposure (input and output) to be able to communicate in the most monolingual-like way with relatively less cognitive load.

In sum, when a monolingual adult begins to acquire a new language later in life, he or she must be able to analyze the L2 input in order to determine where his or her L1 differs on a representational basis, and then, go in search of linguistic information that can satisfy a more target-like L2 representation. The tendency for early L2 speakers to overuse morphosyntactic information from their L1 has to do with the status of this speaker as a very unbalanced bilingual: for any particular semantic plan, the linguistic information instantiated for his or her L1 will always receive the highest activation level for selection even if he or she has begun to acquire the morphophonological information to articulate those features in the L2. As the speaker is able to analyze the L2 input and interact more in the target language, he or she, in principle, has the resources available to develop the same representations as native speakers for all grammatical domains; nevertheless, this isn’t guaranteed as often speakers have difficulty in acquiring domains of grammar that are variable or not highly represented in the input, or where their L1 representation provides sufficient linguistic information for communicative purposes and the
speaker doesn’t perceive a reason to reanalyze it. Finally, these speakers must also be able to learn how to recruit attentional control processes for linguistic purposes so that they can begin to inhibit the activation of their L1 in the expression of their L2. This process may require a high degree of exposure (and interaction) in the target language in order to develop the most robust target-language system possible.

4.3.2.2 Linguistic Outcomes in Adult L2 and for Spanish pronominal systems

Further to the question of how an L2 speaker begins and proceeds through development of a late-acquired language at the early stage, the field of L2 acquisition has also aimed to address the question of ultimate attainment, or rather how complete an adult’s L2 grammar can become (White, 2003; Montrul, 2008). In general, even though the nativist position described above theoretically asserts that adults can acquire the same grammar as children, most evidence from advanced L2 speakers shows some degree of variability in the speakers’ ability to reliably analyze or use functional grammar in a monolingual-like way (Pérez-Leroux & Glass, 1999; Lardiere, 2000; Prévost & White, 2000; Sorace, 2005). In Chapter 3, I discussed the hypotheses associated with this persistent variability in advanced L2 grammars. The proposal most hotly debated and empirically studied today is that which supports the notion of interface vulnerability in contexts of bilingualism (Müller & Hulk, 2001; Sorace 2005, among others). This approach attempts to explain what areas of grammar are most likely to exhibit bilingual effects (are most vulnerable) at the most advanced levels of knowledge/use. Recall that there is some empirical evidence that does not support such a hypothesis (Iverson 2009; Argyri & Sorace, 2007), and as such I suggested that persistent variability in advanced adult grammars is no different than the variability seen in child and developing grammars: that it occurs for domains of grammar where there is some learnability issue in the target language; that is to say where a speaker is not able to clearly figure out from input what the linguistic feature (and distribution of that feature) means from a semantic point of view in the most target-language-specific way. Here, factors such as language pairings, type and quantity of exposure and relative language proficiency (i.e. lexical robustness) can have an impact on how complete the L2 grammar of any one speaker can be (Unsworth, 2008; Montrul, 2008; Tracy, 1995).
In general, therefore, the areas of grammar that have exhibited some persistent variability are those where the syntax-to-semantics/pragmatics mapping does not happen on a one-to-one basis (null subjects, Al-Kasey & Pérez-Leroux, 1998; Pérez-Leroux & Glass, 1999; tense/aspect, Montrul & Slabakova, 2002; Bruhn de Garavito & Valenzuela, 2008; mood, Sánchez-Navarro, 2009; unaccusativity, Sorace, 1993; pronominal morphosyntax, Valenzuela, 2006). For the acquisition of Spanish (or Romance) pronominal morphosyntax by English speakers, most studies have shown that English speakers can acquire the clitic status of unmarked pronominal objects and their basic distributions in simple tensed and untensed clauses (Liceras, 1985; Duffield & White, 1999; Montrul, 2010), but that they may have trouble through the course of acquisition and at advanced stages for properties/features that are more variable or not highly represented in input. For example, Duffield & White (1999) examined object clitic placement in Spanish by native English speakers and found that these speakers had no trouble in acquiring the clitic status of pronouns and their basic distributions, but some advanced speakers had persistent difficulty with clitic-climbing constructions where the clitic position was obligatory (causatives). Nevertheless, English speakers of advanced L2 Spanish performed well in on-line tasks and in judging the grammaticality of the optional clitic-climbing structures under investigation in the current study.

Finally, Montrul (2010) examined English speakers of low-intermediate L2 Spanish, and she too found that these speakers had already acquired the clitic status and basic distribution of object pronouns in Spanish, but had some trouble in the judgment of more complex structures, such as causatives and topicalization structures (clitic-left dislocation, CLLD; also see Valenzuela, 2006 for a study on persistent variability for CLLD with advanced L2 speakers of Spanish). Further, in a narrative production task, low-intermediate L2 speakers of Spanish only produced the unmarked word order (enclisis) for pronominal clitics with complex infinitival predicates even though they appeared to have knowledge of the more marked form (proclisis) where optionality is permitted (Spanish). This was compared with low-intermediate heritage speakers of Spanish who produced the highest rates of proclisis in the narrative task for all groups tested (including monolinguals). In the current study, the aim will be to expand on this investigation and examine if English speakers of advanced L2 Spanish overcome the tendency to resort categorically to the unmarked structure in production, and to what degree they use the same cues (lexical and semantic) as native speakers to deal with the optionality. Further this
study examines the details of optional clitic-climbing more closely by considering the rates of proclisis across a continuum of verb types and the factors which speakers may use to restrict the optionality.

4.4 Chapter Summary

In this chapter I reviewed the factors that contribute to bilingual-like behaviour in contexts of early and late bilingualism, as well as the empirical studies that have examined early and late bilinguals at different stages of development. In general, early bilinguals who are exposed to a heritage or minority language have strong naturalisitic exposure characterized by the vernacular of the target language. In the case of Spanish heritage speakers, few are educated formally in Spanish, and as such, arrive at adulthood with linguistic characteristics typical of this more informal exposure. Conversely, L2 speakers of Spanish in North America typically have patterns of exposure that favour formal language varieties representative of an instructed learning context and writing. Adult L2 speakers often arrive at advanced levels of proficiency with incomplete grammars, especially for grammatical domains that present some learnability challenge (ambiguous or underrepresented input).

While it is not surprising that L2 speakers often arrive at advanced stages of proficiency exhibiting some non-monolingual-like language behaviour, heritage speakers have also been shown to exhibit similar non-monolingual effects at advanced proficiency, despite being native speakers of the target language. The similarity of language outcomes for early and late bilinguals forces us to reanalyze how we view notions of a critical period in L2 acquisition, as well as how we view the notion of complete acquisition for bilingual populations. Previous research has generally shown that late bilinguals can acquire grammatical features not instantiated in their L1 (such as Spanish pronominal clitics and verb incorporation by English speakers). As such, I proceed in the current study by assuming that in principle L2 speakers can achieve the same linguistic outcomes as native speakers (i.e. heritage speakers, monolinguals). The goal of the present study is not to determine if these bilinguals can acquire pronominal clitics or optional clitic-climbing, but rather whether the age of exposure (native vs. non-native) or the context of exposure (monolingual vs. bilingual) to bilingualism has an impact on the way that bilingual speakers deal with the optionality that characterizes Spanish clitic-climbing structures.
5 The Study: Research Questions, Methods and Hypotheses

5.1 Introduction

The aim of this chapter is to outline the methodological components of the current empirical study: the research questions, the profile of participants and the criteria associated with their selection, the experimental tasks and the predictions for how the different experimental groups would perform given prior empirical evidence of the unique linguistic behaviours of early and late bilinguals. As mentioned in Chapter 1, there are three fundamental questions being pursued in this study: 1) questions on categorical structural knowledge; 2) questions on non-categorical lexical knowledge (receptively and productively); and 3) questions on non-categorical semantic knowledge. I will present the research questions, methodology and hypotheses in this chapter on the basis of these three areas of inquiry.

5.2 Research Questions

The questions in this study pertain to bilinguals’ knowledge of clitic-climbing constructions. Spanish speakers must acquire two types of structural knowledge for this grammatical domain. First, they must come to know that a subset of infinitive-selecting verbs of different syntactic classes (functional, raising, control) can undergo the process of verb incorporation when a pronominal object is selected, resulting in a Spanish-specific word order (proclisis).

(71) i. Juan la va a escribir. / Juan ya a escribirla [functional: future tense]
   *John it is going to write / John is going to write it.
   “John is going to write it.”

   ii. Juan la puede escribir / Juan puede escribir la [raising: modal]
      *John it can write / John can write it
      “John can write it.”

   iii. Juan la quiere escribir / Juan quiere escribir la [control: desiderative]
      *John it wants to write / John wants to write it
      “John wants to write it.”

Second, they must come to know that the Spanish-specific structural option (proclisis) is categorically limited when lexical material is embedded between the matrix and infinitive verb of the complex predicate, as in (73):
To examine the structural knowledge of Spanish speakers for this grammatical domain, therefore, I pursue the following questions:

- Do early and late bilinguals of Spanish show evidence (in a receptive task) of instantiating a subset of infinitive-selecting verbs that permit both unmarked (enclisis) and Spanish-specific (proclisis) structural variants similar to monolinguals?

- Do early and late bilingual speakers of Spanish know that embedded negation acts as a categorical constraint blocking the Spanish-specific structural option (verb incorporation, proclisis) available for clitic-climbing constructions?

For the lexical knowledge of optional clitic-climbing in Spanish, Davies (1995) showed that Spanish speakers (monolinguals) organize the optionality presented in (71) above according to cues provided by the matrix verb type of the complex predicate. Here, a main question of inquiry has to do with the lexical organization of the optionality and the limitations imposed on a random grammar:

- Do bilinguals use proclisis as much as monolinguals in production?

- Do early and late bilinguals differ in their use of lexical cues to limit the optionality of Spanish clitic-climbing structures in production (compared to monolinguals and compared to each other)?

- Do Spanish speakers (monolinguals, bilinguals) use lexical cues (matrix verb type) to organize and limit the optionality from a generalized linguistic perspective (in a receptive task)?
Consequently, do monolinguals and bilinguals exhibit a task-based asymmetry on the limitation of grammatical optionality by lexical-based cues provided by the matrix verb of the complex predicate (production vs. receptive tasks)?

Finally, given the widely-held assumption that the relative use of proclisis in optional clitic-climbing structures has to do with the relative semantics (light vs. heavy interpretation) of the matrix verb (Napoli, 1982; Davies 1995), the aim of the current study was to test this notion directly using light and heavy interpretations of the same matrix verb, *ir a* (to be going to), as in (74).

(74) i. Mañana Pedro lo va a comprar. / Mañana Pedro va a comprarlo.  
*Tomorrow Peter it is going to buy. / Tomorrow Peter is going to buy it.  
“Peter is going to buy it tomorrow.” [“light” verb to go: future interpretation]

ii. Pedro la fue a buscar. / Pedro fue a buscarla.  
*Peter it went to look for. / Peter went to look for it.  
“Peter went to look for it.” [“heavy” verb to go: motion interpretation]

Since no previous study has empirically tested the role of matrix verb semantics on the organization and limitation of optionality for clitic-climbing constructions, I pursue the following research questions on the use of interpretational cues to organize and limit optionality in Spanish clitic-climbing constructions:

- Do Spanish monolinguals use interpretational cues to organize and limit grammatical optionality from a generalized linguistic perspective (receptively)?

- Do Spanish-English bilinguals use interpretational cues to organize and limit the grammatical optionality associated with Spanish clitic-climbing?

- How do Spanish-English bilinguals compare against Spanish monolinguals in their use of interpretational cues to limit the grammatical optionality associated with Spanish clitic-climbing from a generalized linguistic perspective (receptively)?

In the following section, I will present the methodology of the current study and then in section 5.4, I will present the hypotheses for the questions above.
5.3 Methodology

5.3.1 Participants

Forty (n=40) Spanish speakers participated in this study, including early adult Spanish-English bilinguals (n=10) of advanced Spanish proficiency, late adult Spanish-English (L2) bilinguals (n=16) of advanced Spanish proficiency and Spanish native speakers with very low English proficiency (control group) (n=14). I recruited participants in Toronto, Ontario in university environments and at a variety of ESL schools; all participants were either pursuing some form of post-secondary education or had completed this prior to participating in the study.

All participants read and filled out a consent form before completing the experimental tasks. During the testing session, all participants also completed a language history interview based on guidelines set out in an in-depth language history questionnaire (see Appendix B). These guidelines included information related to age of onset of bilingualism, mode of language learning, language of formal education, self-ratings of proficiency in Spanish and English (and other languages), experience with language immersion (i.e. stays in Spanish-/English-speaking communities, etc.), language used at work and home, parental language history, etc.

Bilingual groups also completed a narrative task in which they were asked to narrate a wordless picture book, *A boy, a dog and a frog*. All bilingual participants provided the narration in Spanish in order to verify language proficiency, and adult heritage bilinguals also provided the narration in English at the end of the session to determine patterns of dominance among the heritage bilinguals. Late adult L2 bilinguals were not required to complete the narrative task in English because they were all native speakers of English (monolingual L1 development). However, three participants completed the narrative task in English in order to provide a baseline of English proficiency for bilinguals. Three additional monolingual speakers of English and four Spanish native speakers from the control group completed the narrative task in English and Spanish respectively in order to provide monolingual samples of the narrative task.

Participants for the Spanish control group consisted of 14 Spanish native speakers who had spent less than one year of cumulative time in an English-language community. Their exposure to English had been limited to this short stay in an English-language community and some classroom instruction in their home country. They all rated themselves with poor
proficiency in English (or other languages to which they had been exposed). The native speakers came from a variety of Spanish-speaking countries including Mexico (n=9), Venezuela (n=1), Spain (n=3) and Cuba (n=1). Given that Davies (1995) showed that dialectal variation does not influence the variability to be tested in this study, no participants were excluded on the basis of the dialect to which they had been exposed.

The late adult L2 bilingual group consisted of 16 native English speakers whose first contact with Spanish had occurred after 15 years of age in an instructional setting. To qualify for participation in the study, participants must have initially learned Spanish in a classroom setting and either concurrently or subsequently spent a minimum of 12 months cumulative time in a Spanish-speaking country. Finally, participants were included in the study only if they both rated themselves as highly proficient in Spanish (general language ability) and were able to successfully (and with ease) complete both production and narrative tasks in Spanish.

While a higher number of early adult heritage bilinguals were considered for this study (n=25), only ten of those were selected for analysis based on fulfilling two criteria. First, a linguistic criterion whereby speakers a) had to identify themselves as highly proficient in both Spanish and English; and b) were successfully able to complete production and narrative tasks in Spanish (with ease). Second, an exposure criterion, whereby a) they identified that they had been exposed to both Spanish and English before the age of 2;0; b) they had either been born in or resided in an English language community from at least 2;0, and consequently, been formally educated (primary and secondary education) in English; c) they had native Spanish-speaking parents who immigrated to English language communities as adults; d) they identified that they had relatively consistent exposure (and fluency) in both languages from childhood to adulthood; and e) they had been exposed to at least one semester of classroom learning in Spanish.

Since the majority of all bilingual participants had been born and formally educated in Ontario, most revealed that they had some exposure to French as a second language (a core component of provincial curriculum in Ontario from grade 4 to high school). In order to be included in this study, participants had to identify that they were not more than moderately fluent in their general French language abilities, and importantly, that they were more highly fluent in Spanish. In standard accounts, French is not a restructuring language (does not employ verb
incorporation) and as such, those with more advanced knowledge of French would have an additional linguistic resource supporting the unmarked structure of complex infinitival predicates. Seeking out participants with no French background was impractical given the population available for testing; as such, the aim here was to find those with the least amount of exposure to French and those who self-identified as poor speakers of the language.

Finally the Spanish dialect to which all bilingual participants had been exposed ranged widely, including varieties of Mexico, Central America, South America, the Caribbean and Spain. Recall that Davies (1995) examined region as an independent variable in his corpus study and found that the relative distribution of pronominal clitics with complex predicates did not vary according to regional variety. Consequently, similar to monolinguals, bilingual participants were not excluded on the basis of the Spanish variety to which they had been exposed.

Table 5.1 summarizes the criteria used to determine the suitability of participants for the current study:
Table 5.1

Criteria for selection of participants in the current study

<table>
<thead>
<tr>
<th></th>
<th>Spanish monolinguals (L1 control group)</th>
<th>Early adult heritage L2 bilinguals</th>
<th>Late adult L2 bilinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linguistic criteria</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Self-ratings</em></td>
<td>Spanish: native</td>
<td>Spanish: advanced/native</td>
<td>Spanish: advanced/native-like</td>
</tr>
<tr>
<td></td>
<td>English: poor</td>
<td>English: native</td>
<td>English: native</td>
</tr>
<tr>
<td><strong>Communicative fluency in Spanish</strong></td>
<td>Native</td>
<td>Highly fluent (ability to easily complete production tasks)</td>
<td>Highly fluent (ability to easily complete production tasks)</td>
</tr>
<tr>
<td><strong>Exposure criteria</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Place of birth/infancy</em></td>
<td>Spanish-speaking country (no dialectal restrictions)</td>
<td>English-speaking country OR arrival before 2;0</td>
<td>English-speaking country</td>
</tr>
<tr>
<td><em>Early language exposure &lt;2;0</em></td>
<td>Spanish</td>
<td>Spanish and English</td>
<td>English</td>
</tr>
<tr>
<td><em>Language of formal education (to high school)</em></td>
<td>Spanish</td>
<td>English</td>
<td>English</td>
</tr>
<tr>
<td><em>1st exposure to Spanish-English bilingualism</em></td>
<td>Limited</td>
<td>&lt; 2;0</td>
<td>&gt; 15 years old</td>
</tr>
<tr>
<td><strong>Mode of Spanish learning</strong></td>
<td>Naturalistic (L1) Formal education</td>
<td>Naturalistic (L1) Limited classroom</td>
<td>Classroom Naturalistic (L1 &amp; L2)</td>
</tr>
<tr>
<td><em>Language Immersion</em></td>
<td>&lt; 1 year (cumulative) English-speaking country</td>
<td>minimum 6 months (cumulative) Spanish-speaking country</td>
<td>minimum 1 year (cumulative) Spanish-speaking country</td>
</tr>
<tr>
<td><em>Other L2 (eg, French)</em></td>
<td>None</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Parental L1</em></td>
<td>Spanish</td>
<td>Spanish</td>
<td>English</td>
</tr>
<tr>
<td><em>Parental Place of birth &amp; formal education (to high school)</em></td>
<td>Spanish-speaking country</td>
<td>Spanish-speaking country</td>
<td>English-speaking country</td>
</tr>
</tbody>
</table>

Finally, in order to assess the relative Spanish fluency of bilingual speakers, the narrative data was assessed independently by two native Spanish-speaking judges (see Cuza 2008). No standardized proficiency test was administered in the testing session in order to avoid testing fatigue. The narratives of all bilinguals (n=26) and Spanish control participants (n=4) were presented randomly to judges who were asked to score the speakers’ proficiency in relation to syntactic, lexical (vocabulary), phonological (pronunciation), and overall fluency and grammar on a scale of 1 to 5, with 5 being reserved for the most proficient speakers. The same assessment of the English narratives (bilinguals n=13; English monolinguals n=3) was carried out.
independently by two native English-speaking judges. A summary of the participants’ profiles is presented in Appendix C.

5.3.2 Structures under Analysis

As shown in previous chapters, Spanish and English overlap in the surface representation of word order for pronominal objects with infinitival constructions. English instantiates invariably post-verbal object pronouns while Spanish optionally instantiates both post-verbal pronominal objects (enclisis) and pre-verbal pronominal objects (proclisis) for a variety of infinitival constructions. Despite apparently optional word orders, however, I also showed that the apparent word order variability is better analyzed under Minimalist terms as variability in the order of Merge operations that the matrix verb, main infinitival verb and pronominal argument undergo in building the infinitival construction. In English and Spanish the pronominal object (DP in English and pro in Spanish) merges first to the infinitive and then this simple predicate merges to the tensed matrix/auxiliary verb of the expression, resulting in a surface structure with post-verbal pronouns (enclisis for Spanish). In English only this order of merge operations is permitted; however, in Spanish an alternative structure-building option is available if a pronominal argument is selected: the verbs undergo the process of verb incorporation (to form a complex predicate, in Masullo’s terms) prior to entering syntax when the merge of the pronominal argument (pro) occurs; pronominal objects subsequently surface in the pre-verbal position (proclisis) as the result of morphophonological processes related the tense and complexity of the predicate. A variety of factors, such as lexical, semantic and structural constraints may act to limit the optionality that occurs when a pronominal argument is selected in Spanish. The aim of the experiments, therefore, was to assess the degree to which a bilingual speaker can acquire the subtle constraints (both categorical and non-categorical) that limit the degree to which speakers apply verb incorporation for this grammatical domain.

First, I evaluated the knowledge of lexical constraints through minimal pairs of sentences employing both pre- and post-verbal pronouns (proclisis and enclisis respectively) constructed with a variety of matrix (auxiliary-like) verbs, including ir a ‘to be going to’, poder ‘to be able to’, querer ‘to want to’, necesitar ‘to need to’, preferir ‘to prefer to’ and evitar ‘to avoid’. I selected these particular matrix verbs to represent the continuum of variability that has been shown to occur for this grammatical domain with monolingual Spanish speakers (Davies, 1995;
Strozer, 1976): highest rates of proclisis with *ir a* (86%) with declining rates thereafter, *poder* (60%), *querer* (47%), *necesitar* (33%), *preferir* (15%), *evitar* (0%). Second, I evaluated the effect of predicate semantics on the formation of infinitival constructions with minimal pairs (proclisis and enclisis) of two types of sentences employing the verb *ir a* ‘to be going to’: first, sentences in the present tense and marked by an adverb pointing to future time for a more grammaticalized, future interpretation of the event (and matrix verb); and second, sentences in the perfective past tense for a more literal motion interpretation of the event (and matrix verb). Finally, in order to evaluate the knowledge that Spanish speakers had of the structural conditions that limit the variability described above, I presented minimal pairs of sentences with both pre- and post-verbal pronouns which employed embedded negation between the matrix and infinitival verb of the construction. Unlike the other conditions, which mostly represented non-categorical constraints of the variability (except, according to Strozer, *evitar*), the structural condition represented a categorical constraint limiting the variability, and as such, I expected the speakers who had full knowledge of the constraints that underlie this grammatical domain to accept only the unmarked variant (enclisis). Table 5.2 summarizes the conditions described above:

Table 5.2

*Conditions used to test Verb Incorporation in Spanish*

<table>
<thead>
<tr>
<th>Lexical variable</th>
<th>Juan la va a comprar.</th>
<th>[PRO √√]</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ir a</em> (future—present</td>
<td>Juan va a comprarla.</td>
<td>[ENCL √]</td>
</tr>
<tr>
<td>“to be going to”</td>
<td>John is going to buy it.</td>
<td></td>
</tr>
<tr>
<td><em>poder</em></td>
<td>Juan la puede comprar.</td>
<td>[PRO √]</td>
</tr>
<tr>
<td>“to be able to”</td>
<td>Juan puede comprarla.</td>
<td>[ENCL √]</td>
</tr>
<tr>
<td><em>querer</em></td>
<td>Juan la quiere comprar.</td>
<td>[PRO √]</td>
</tr>
<tr>
<td>“to want to”</td>
<td>Juan quiere comprarla.</td>
<td>[ENCL √]</td>
</tr>
<tr>
<td><em>necesitar</em></td>
<td>Juan la necesita comprar.</td>
<td>[PRO √]</td>
</tr>
<tr>
<td>“to need to”</td>
<td>Juan necesita comprarla.</td>
<td>[ENCL √√]</td>
</tr>
<tr>
<td><em>preferir</em></td>
<td>Juan la prefiere comprar.</td>
<td>[PRO √]</td>
</tr>
<tr>
<td>“to prefer to”</td>
<td>Juan prefiere comprarla.</td>
<td>[ENCL √√]</td>
</tr>
<tr>
<td><em>evitar</em></td>
<td>*/?Juan la evitó comprar.</td>
<td>[PRO*]</td>
</tr>
<tr>
<td>“to avoid”</td>
<td>Juan evitó comprarla.</td>
<td>[ENCL √]</td>
</tr>
</tbody>
</table>

John avoided buying it.
I used two types of tasks to test these conditions presented in Table 5.2. First, a production task tested the lexical condition in order to evaluate the choice that speakers make in the process of structure-building for infinitival expressions that employ three different matrix verbs, *ir a, querer* and *preferir*. Second, an acceptability-preference task tested the knowledge that speakers have of the lexical, semantic and structural constraints available to limit the optionality characteristic of this grammatical domain. I could subsequently compare the results from the production task directly against those for the same verbs in the acceptability-preference task in order to determine if bilingual speakers use lexical cues both in production and in the organization of the grammar from a generalized perspective (in a receptive task), or if lexical cues only operate when the speaker has to make a choice between competing structural options (unmarked vs. Spanish-specific) for expressive purposes.

In the production task, there were a total of fifteen test tokens (5 test tokens x 3 verb types) in a semi-randomized order with seven distracters. The distracter tokens for the production task focused on eliciting information about non-argumental components of an event (eg., *what colour?*, *how many?*, etc.). For the acceptability-preference task, there were a total of 32 test tokens (4 test tokens x 8 conditions) in a semi-randomized order with 15 distracters. The distracter tokens for the acceptability-preference task tested three grammatical areas (5 tokens x 3 types): 1) preterite-imperfect distinction; 2) placement of adverbs in perfect tense predicates; and 3) transitive/intransitive causative constructions (*Se derramó la leche* ‘the milk spilled’ vs. *Derramé la leche* ‘I spilled the milk’). There were an additional four training tokens at the
beginning of the test presenting contexts and minimal pairs related to negation, adverbial placement and grammatical aspect.

5.3.3 Tasks and Testing Session

There were four main components of this study: i) a Picture Elicitation Task (PET); ii) an Acceptability-Preference Task (APT); iii) a Narrative Task (A boy, a dog and a frog); and iv) a Language History Interview. During the testing session (and in prior communication with participants), I only used Spanish to promote the most monolingual mode of communication possible (Grosjean, 1998). Since the testing occurred at the University of Toronto, it was difficult to control strictly for mode of communication because the participants most certainly identified the tester, a graduate student at an English-language university, as a bilingual speaker of English and Spanish. Nevertheless, an effort was made to communicate with participants strictly in Spanish, both during the testing session and in communication prior to the session. Consequently, all instructions, documents (consent form) and dialogue between the tester and the participant was carried out in Spanish.

The testing session took an average of one hour to complete and was broken down into six blocks: 1) Consent form; 2) Picture Elicitation Task; 3) Acceptability-Preference Task: Part I; 4) Narration of A boy, A dog and a frog; 5) Language History Interview; and 6) Acceptability-Preference Task: Part II. The APT was broken into two blocks because the high number of tokens involved in this task (52) may have led to participant fatigue. I asked all heritage speakers and three L2 speakers to narrate the same frog story in English at the end of the session.

In the following sections I will present the experimental tasks in detail in relation to the goals of the methodology and the procedures involved in their administration. In Appendices E, F, G and H, I have included the instructions for both tasks, the full response sheet for the APT and the complete contexts and target sentences/questions heard by participants for both tasks.

5.3.3.1 Picture Elicitation Task (PET)

The goal of the PET was to investigate bilingual speaker preferences for the application of optional structure-building processes for Spanish infinitival constructions: a Spanish-specific structure-building process that results in pre-verbal pronouns (proclisis) or an unmarked
structure-building process that overlaps with English and results in post-verbal pronouns in Spanish (enclisis). Further to this goal, this task also aimed to establish whether bilingual speakers were sensitive to the lexical factors that have been shown to limit the variability that Spanish monolinguals apply for this grammatical domain in spontaneous speech (Davies 1995).

The PET has been frequently used in child studies to test knowledge of binding principles and pronominal morphosyntax in the course of language acquisition (Jacubowicz, 1983, Schaeffer, 1997, cf. Pérez-Leroux, Pirvulescu & Roberge, 2006). It was an ideal method, therefore, to investigate how fluent speakers of Spanish apply structure-building processes that include pronominal clitics. Nevertheless, due to the complex nature of the verbal predicate involved, some modifications of the procedures used in the above studies had to be made in order to get participants to use complete SVV\textsubscript{inf}O or SOVV\textsubscript{inf} sentences.

In each picture, contrasting characters were depicted who were shown to be imagining what they ‘were going to do’, ‘wanted to do’ or ‘preferred to do’ with the same object. Three pairs of characters were depicted continuous with a specific matrix verb: i) Marta and Pedro, pre-teenage siblings were matched with the matrix verb \textit{ir a} ‘to be going to’; ii) María and Anita, young twin sisters were matched with the matrix verb \textit{querer} ‘to want to’; and iii) Jorge and Pilar, a newly married couple were matched with the matrix verb \textit{preferir} ‘to prefer to’. Participants were shown each picture and were read a short context by the tester, which always ended in a question that introduced the target predicate and direct object as a DP. A sample token with context, question (both aural), picture and the ideal target answer is presented in (75).
Las gemelas están sentadas en la sala de estar. La ventana está medio abierta y Anita tiene calor pero María tiene frío. ¿Qué quiere hacer cada niña con la ventana?

The twins are sitting in the living room. The window is half open and Anita is hot, but Maria is cold. What does each girl want to do with the window?

b) Ideal target answer: Anita quiere abrir la y María quiere cerrarla.

Anita wants to open it and Maria wants to close it.

There were four possible target structures (counted) for each full clause:

1) Anita la quiere abrir (proclisis)
   Anita it wants to open

2) Anita quiere abrirla (enclisis)
   Anita wants to open it

3) Anita quiere abrir la ventana (object DP)
   Anita wants to open the window

4) Anita quiere abrir Ø. (object omission)
   Anita wants to open Ø.

In the case of sentences 2) and 3), the unmarked structure-building process would have been applied by speakers, namely merge first of the argument and the infinitive and then subsequent merge of the simple predicate to the matrix (tensed) verb. Conversely, in the case of the first
sentence, speakers would have been using the more Spanish-specific variant for structure-building (verb incorporation): lexical incorporation of the matrix and infinitival verb (forming a complex predicate) and then merge of the pronominal clitic. For the cases of object omission, it is uncertain which process of structure building would have been used; I discuss these cases more in Chapter 7.

Given that each picture depicted two characters acting in different ways upon a direct object, there were a total of ten possible test tokens per verb: 5 pictures per verb x 2 sentences per picture. Three training items testing non-argumental knowledge introduced the three pairs of characters, and the test items were semi-randomized among seven distracters, which also depicted the three pairs of characters.

Prior to beginning the test, I instructed participants to listen closely to the context and question presented and respond in a manner that most closely resembled the question. I made this instruction to ensure maximal use of the target matrix verb (ir a, querer and preferir) for each image. I further instructed the participants that they could ask me to repeat the context and/or the question, but in no case did this occur. The answers of participants were digitally recorded and later transcribed into a database (SPSS) for statistical analysis.

5.3.3.2 Acceptability-Preference Task (APT)

The APT was a unique design that took into account methodologies used to elicit categorical and non-categorical judgments related to minimal pairs (Tarone et al., 1994). The aim of the task was to test the knowledge that speakers have of both categorical limitations on pronominal placement with infinitival constructions as well as the tendencies that characterize speaker preferences when limitations are not categorical, and so, the degree of variability permitted in any type of grammar (e.g., monolingual, bilingual).

Given that the aim of this task was to test the underlying knowledge that speakers have of the limiting factors that characterize the variability of the grammatical domain under investigation, the APT was designed so that participants only had contact with aural stimuli related to target sentences. Showing target sentences in written form may have led to confounding results for two reasons. Davies (1995) showed that monolingual speakers of Spanish tend to use more post-verbal pronouns overall in written communication as compared to oral
communication, and therefore, seeing the sentences may have confounded preferences related to the two modalities. Further, late learners of Spanish tend to have extensive experience with written forms of the language in an instructional setting; consequently, in seeing a written form of the target sentences, late bilinguals may revert to the declarative knowledge they have for this grammatical domain (e.g., what their teacher said about the structures), instead of accessing the procedural knowledge they may have acquired on the basis of more naturalistic input.

The task, therefore, required that participants listen to contexts and target sentences (minimal pair) that had been previously recorded by a native speaker of Colombian Spanish (non-coastal variety), and mark their preferences on a testing sheet provided to them. The contexts for each scenario were constructed so as to avoid the use of infinitival predicates or pronominal clitics in order to avoid any priming effect that could occur in the assessment of the minimal pairs.

The testing sheet (see Appendix E) was enumerated and provided information related to the agent of the event and one other guiding detail (such as, direct object). I included this information to ensure that the participant was marking the correct number for each recorded context. Below this information were two boxes, ‘1st is acceptable’ and ‘2nd is acceptable’. I instructed participants (in Spanish) to listen to the target sentences that followed the context and independently decide if the 1st and 2nd were both acceptable sentences of Spanish or if only 1st or 2nd was acceptable. If the participant marked that both 1st and 2nd were acceptable, then they would continue on to the second part of the question where they would indicate their preferences. The leading question was, ‘If both are acceptable and you were to say this sentence, how would you say it?’ Here they could mark ‘first’, ‘second’, or ‘I can’t decide’. I instructed them to mark ‘I can’t decide’ if they were ambivalent about which structure was better for the context. A sample of what the participants would see on the test sheet is illustrated below in (76b) and the complete testing sheet (version A) is presented in Appendix G. A sample test question, including the aural prompt and the test sheet, is provided in (76).
(76) a) Context and target sentences (in italics) heard via recording:

*El año pasado Carlos compró un carro y recientemente se quedó sin empleo.*

“Last year Carlos bought a car and recently he lost his job”

Ahora Carlos *necesita venderlo*.

(Now Carlos *needs* to sell *it*)

Ahora Carlos *lo necesita vender*

(Now Carlos *it needs* to sell)

b) Testing sheet (visual stimulus):

1. Carlos y el carro
   
   □-1ª es aceptable □-2ª es aceptable

   Si las dos son aceptables, y si tú fueras a decir esto, ¿cómo lo dirías?:
   
   ___ La primera
   ___ La segunda
   ___ No puedo decidir

   Four training contexts represented four combinations of answers: only 1ª was acceptable, only 2ª was acceptable, both were acceptable with a preference for the first and both were acceptable with a very subtle change that was indiscernible to most participants (thus, leading most of them to mark ‘I can’t decide’). I explained the process in detail during the training component so that by the time participants began the test, they felt comfortable with the methodology and appeared to be answering in ways that reflected their understanding of the task.

   There were four different versions of the acceptability-preference task. In all versions, minimal pairs of test sentences (with proclisis and enclisis) were presented in a 2x2 order: two of each condition with pre-verbal pronouns first and two of each condition with post-verbal pronouns first. Version A of the protocol included the same order of semi-randomized test items and distracters with Version A1 and A2 accounting for a counter-balanced presentation of minimal pairs. Version B was a different semi-randomized order of test items and distracters (e.g., the last tokens from Version A appearing first, etc) with Version B1 and B2 accounting for a counter-balanced presentation of minimal pairs of test sentences. The order of training items was consistent across all versions of the test and the order of minimal pairs was not counter-balanced.
Overall, in judging the acceptability and/or preference of the two word orders, proclisis and enclisis, for the six verb types being examined in the Acceptability-Preference Task (APT)—ir a (future), poder, querer, necesitar, preferir, evitar—or in producing the two word orders with ir a, querer and preferir in the Picture Elicitation Task (PET), participants were not expected to categorically select one word order over the other. With the exception of the verb evitan (see Strozer 1976), all verbs examined for this constraint have been shown to permit some degree of variability in the grammars of Spanish native speakers (Myhill, 1988, Davies, 1995). As such, in the APT, I expected that if participants included a particular verb in the subset of infinitive-selecting verbs that permits verb incorporation, they would exhibit relatively high rates of acceptability of both the proclisis and enclisis versions of the sentences presented to them (well over 50%). As such, the objective here was to explore the factors that Spanish speakers use to restrict the core optionality for this grammatical domain by testing the relative acceptability/preference and production of proclisis.

5.4 Hypotheses of Study

In section 5.2 above, I outlined three main groups of research questions related to the structural, lexical and semantic variables being tested in this study. Here, I will consider how I expected the three groups of Spanish speakers being tested in this study to perform on the tasks outlined in section 5.3.3. In general, there are three basic hypotheses for any independent variable:

1) No group differences (null hypothesis).

2) Bilingual effect hypothesis (L1 ≠ 2L1, L2)—context of exposure (single- vs. dual-language context of exposure) appears to have a greater impact on the way that Spanish speakers organize and limit grammatical optionality for Spanish clitic-climbing constructions.

3) Age effect hypothesis (L1, 2L1 ≠ L2)—age of exposure (early vs. late exposure) appears to have a greater impact on the way that Spanish speakers organize and limit grammatical optionality.

First, as outlined in section 5.2, there are two questions regarding structural knowledge that I am pursuing in this study:
• Do early and late bilinguals of Spanish show evidence (in a receptive task) of instantiating a subset of infinitive-selecting verbs that permit both unmarked (enclisis) and Spanish-specific (proclisis) structural variants similar to monolinguals?

• Do early and late bilingual speakers of Spanish know that embedded negation acts as a categorical constraint blocking verb incorporation and proclisis?

In the first case, I expect the null hypothesis to hold: given evidence that we have for the ability of early and late Spanish-English bilinguals to acquire the Spanish-specific structural variant for optional clitic-climbing constructions (Liceras, 1985; Duffield & White, 1999; Montrul, 2004, 2010), all groups of Spanish speakers here should show evidence that they have acquired the necessary structural knowledge (verb incorporation) that results in a Spanish-specific word order (proclisis) for this domain. In the second case, even though pronominal complex predicates with embedded negation are not likely to be a construction that Spanish speakers have a great deal of experience with in input, Spanish speakers should be able to acquire the categorical constraint associated with embedded negation as long as they have knowledge of the universal constraints related to verb incorporation. If we assume a full access model of late acquisition, such as the Full Transfer Full Access model of Schwartz and Sprouse (1996), then all groups should perform similarly in identifying the difference between the availability of proclisis with regular complex predicates (lo + querer/poder + infinitive) as compared to those with embedded negation (*lo + querer/poder + no + infinitive).

Further, regarding the lexical cues that Spanish monolinguals have been shown to use (in production) to limit grammatical optionality in clitic-climbing constructions (Davies (1995), I have outlined two questions related to the way that bilinguals use proclisis:

• Do bilinguals use proclisis as much as monolinguals in production?

• Do early and late bilinguals differ in their use of lexical cues to limit the optionality of Spanish clitic-climbing structures in production (compared to monolinguals and compared to each other)?

In the first case, I expect that bilinguals will use proclisis at overall lower rates than monolinguals due to either or both processing and cross-language effects; consequently, I expect the bilingual
groups to exhibit more of an enclisis bias in production than monolinguals, supporting the bilingual effect hypothesis. In the second case, I expect the age of exposure to bilingualism to have an impact on the degree to which bilingual speakers will make use of lexical cues in production to limit the use of proclisis, thus supporting the age effect hypothesis. Even if L2 speakers exhibit some sensitivity to verb type in the use of proclisis, I expect early bilinguals to be able to overcome psycholinguistic/cross-language effects that lead to an enclisis bias more than late bilinguals (leading to a possible interaction effect here). As such, I expect early bilinguals to outperform late bilinguals on their use of lexical cues to limit the optionality: the highest rates of proclisis with *ir a* (to be going to), more moderate rates of proclisis with *querer* (to want to), and low rates of proclisis with *preferir* (to prefer to).

To explore the production biases of Spanish speakers for this domain, I tested the speakers’ sensitivity to lexical cues in a receptive task. The following questions were explored for the receptive performance of the participants in this study, as well as for a comparison in the way that these participants use lexical cues to limit optionality in receptive and production tasks.

- Do Spanish speakers (monolinguals, bilinguals) use lexical cues (matrix verb type) to organize and limit the optionality from a generalized linguistic perspective (in a receptive task)?

- Consequently, do monolinguals and bilinguals exhibit a task-based asymmetry on the limitation of grammatical optionality by lexical-based cues provided by the matrix verb of the complex predicate (production vs. receptive tasks)?

Similar to the expectations for the production task, all groups were expected to exhibit some sensitivity to lexical cues in organizing and limiting proclisis from a general linguistic perspective; however, again I expect age of exposure to bilingualism to have an impact on the degree to which bilingual speakers use these cues, supporting the age effect hypothesis. As such, I expect early bilinguals to outperform late bilinguals here, exhibiting more of a pronounced difference between the acceptability of proclisis across different verb types, leading to an expected interaction effect. I expect late bilinguals to underperform in this regard due to the strong experience that they have in using the unmarked structure (enclisis) overall (also seen
empirically in Montrul, 2010), and possibly due to indirect cross-language effects provided by their L1.

Consequently, I expect L2 speakers to exhibit the strongest task-based asymmetry on the relative acceptance and production of proclisis by lexical type: while accepting proclisis at relatively high (and consistent) rates across lexical types, they will show the strongest enclisis bias in production of all groups. Given that I also expect heritage speakers to exhibit more of an enclisis bias than monolinguals in production, then I expect that the bilingual effect hypothesis will account for limitations of proclisis that bilingual groups make by task type.

Finally, I have outlined three questions for the examination of semantics on the organization and limitation of optionality for Spanish clitic-climbing:

- Do Spanish monolinguals use interpretational cues to organize and limit grammatical optionality from a generalized linguistic perspective (receptively)?

- Do Spanish-English bilinguals use interpretational cues to organize and limit the grammatical optionality associated with Spanish clitic-climbing in a receptive task?

- How do Spanish-English bilinguals compare against Spanish monolinguals in their use of interpretational cues to limit the grammatical optionality associated with Spanish clitic-climbing from a generalized linguistic perspective (receptively)?

Based on the assumptions (Napoli, 1982) that verb incorporation is most likely to occur when a matrix verb (like ir a (to be going to) has a light functional interpretation (future tense) as compared to a heavy lexical interpretation (motion), I expect that monolingual speakers of Spanish will use interpretational cues to organize and limit grammatical optionality for this domain. As such, even if they still accept high rates of proclisis for the verb ir a (as a lexical choice), they will exhibit significant differences in the degree to which proclisis is acceptable/prefereed on the basis of the relative interpretation (future vs. motion) of the event.

For the second and third questions above, I predict that an age effect will emerge. Given the early and mostly naturalistic, native exposure associated with contexts of heritage bilingualism, I expect heritage speakers to perform most like monolinguals on their use of tense/interpretational cues in limiting the optionality for this grammatical domain. L2 speakers, on the other hand, have
more varied experiences with exposure (native/non-native; aural/visual input, etc.) and have been shown to be more successful in acquiring lexically-based distinctions of their L2 (e.g., gender, Unsworth, 2008) than in acquiring semantic/pragmatic-related distinctions (e.g., aspect, mood, Slabakova, 2008; Montrul, 2008). As such, I expect that L2 speakers may not attest to semantic cues to limit this optionality as well as native speakers.

In Table 5.3 I have summarized the independent variables being tested along with the hypotheses outlined above.

Table 5.3

*Hypotheses for structural, lexical and semantic variables for the current study*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural</td>
<td>1) Verb incorporation, proclisis -subset functional/raising/control verbs</td>
<td>1) no group differences</td>
</tr>
<tr>
<td></td>
<td>2) Embedded negation</td>
<td>2) no group differences</td>
</tr>
<tr>
<td>Lexical</td>
<td>1) Production task</td>
<td>1) i. Bilingual effect (L1 ≠ 2L1, L2)</td>
</tr>
<tr>
<td></td>
<td>i. overall use of proclisis</td>
<td>1) ii. Age effect (L1, 2L1 ≠ L2)</td>
</tr>
<tr>
<td></td>
<td>ii. use lexical cues to limit proclisis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2) Receptive task</td>
<td>2) Age Effect (L1, 2L1 ≠ L2)</td>
</tr>
<tr>
<td></td>
<td>-use lexical cues to limit proclisis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3) Receptive vs. Production task</td>
<td>3) Bilingual effect (L1 ≠ 2L1, L2)</td>
</tr>
<tr>
<td></td>
<td>-task-based asymmetry</td>
<td></td>
</tr>
<tr>
<td>Semantic</td>
<td>1) Interpretation of <em>ir a</em> (future vs. motion)</td>
<td>1) Age Effect (L1, 2L1 ≠ L2)</td>
</tr>
</tbody>
</table>

In Chapter 6, I present results of the current study on the basis of the hypotheses outlined above, and then in Chapter 7 I review the empirical results in relation to the macro-variable being examined in this study, age of exposure to bilingualism.
6 The Study: Results

In this chapter I present the results of the empirical study outlined in chapter 5. I present these results according to the three main independent variables being tested in this study: 1) syntactic knowledge of verb incorporation; 2) knowledge of lexical factors that limit verb incorporation; 3) knowledge of semantic factors that limit verb incorporation. Here the results are presented for proclisis given that all participants judged enclisis (the unmarked word order) as an acceptable word order variant for all the structures tested at rates above 90% (an expected outcome).

6.1 Syntactic Knowledge

I pursued two questions related to the syntactic knowledge that early and late bilinguals would exhibit for Spanish clitic-climbing constructions:

1) Do early and late bilinguals of Spanish show evidence having acquired the process of verb incorporation (and thus, proclisis) for a subset of infinitive-selecting (similar to monolinguals)?

2) Do early and late bilingual speakers of Spanish know that embedded negation acts as a categorical constraint blocking the Spanish-specific structural option (verb incorporation, proclisis) available for clitic-climbing constructions?

Given previous studies on Spanish clitic-climbing (Liceras, 1985; Duffield & White, 1999; Montrul, 2010), and assuming a full access model of late acquisition (i.e. Full Transfer Full Access, Schwartz & Sprouse, 1996), I predicted that all groups would perform similarly in exhibiting knowledge of verb incorporation (and thus proclisis) across a range of infinitive-selecting verbs, and exhibit knowledge of the way that embedded negation acts as a categorical constraint blocking the availability of proclisis for this domain.

I examined the two questions above using the Acceptability-Preference task (APT) described in Chapter 5. I present the results for the first question—whether bilinguals are able to acquire and apply the process of verb incorporation across a variety of verb types—in Table 6.1.
Table 6.1

Acceptability of proclisis by group and verb type

<table>
<thead>
<tr>
<th></th>
<th>Ir a (fut)</th>
<th>Poder</th>
<th>Querer</th>
<th>Necesitar</th>
<th>Preferir</th>
<th>Evitar</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 speakers</td>
<td>96.9%</td>
<td>98.4%</td>
<td>96.9%</td>
<td>90.6%</td>
<td>87.5%</td>
<td>89.1%</td>
</tr>
<tr>
<td>%SD</td>
<td>12.50</td>
<td>6.25</td>
<td>8.54</td>
<td>22.13</td>
<td>22.36</td>
<td>25.77</td>
</tr>
<tr>
<td>Heritage</td>
<td>97.5%</td>
<td>97.5%</td>
<td>95.0%</td>
<td>77.5%</td>
<td>82.5%</td>
<td>75.0%</td>
</tr>
<tr>
<td>%SD</td>
<td>7.91</td>
<td>7.91</td>
<td>10.54</td>
<td>36.23</td>
<td>31.29</td>
<td>42.49</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>91.1%</td>
<td>80.4%</td>
<td>85.7%</td>
<td>66.1%</td>
<td>71.4%</td>
<td>55.4%</td>
</tr>
<tr>
<td>%SD</td>
<td>12.43</td>
<td>24.37</td>
<td>18.90</td>
<td>31.94</td>
<td>32.31</td>
<td>32.79</td>
</tr>
</tbody>
</table>

On average, speakers accepted proclisis as an acceptable word order variant for all verbs tested at rates around 70% or more. Only monolinguals failed to accept proclisis with the verb *evitar* (to avoid) at rates above chance (55%). Although not completely ill-formed from a structural point of view, the use of proclisis with the verb *evitar* appears to sound quite odd for monolingual Spanish speakers. These results generally coincide with Strozer’s suggestion that *evitar* is a non-restructuring verb, unlike the other five verbs tested in this task. Heritage speakers and L2 speakers, on the other hand, generally accept *evitar* as a restructuring verb.

Further, while monolinguals accept proclisis with *necesitar* only two-thirds of the time, L2 speakers accept proclisis with *necesitar* at rates above 90%. Here, the tendency for L2 speakers to assign relatively equal status to proclisis for *ir a, poder, querer* and *necesitar* (>90%) may be the result of residual declarative knowledge that these speakers acquired in a classroom setting. Nevertheless, when we examine these results on the basis of speaker preferences, we will see that L2 speakers exhibit stronger tendencies to differentiate these verb types than is evident in the acceptability portion of the task.

In sum, here the predictions of the first question outlined for the structural variable are confirmed: there are no major group differences in knowledge of the Spanish-specific syntactic process (verb incorporation) that results in the surfacing of a pronominal clitic in the pre-verbal position (proclisis) for Spanish infinitival constructions. All groups appear to accept all verbs as being part of a subset of clitic-climbing verbs, although monolinguals don’t appear to include the verb *evitar* in this subset as consistently as bilinguals. Overall L2 speakers exhibit the most consistent rates of acceptability of proclisis (and thus, verb incorporation) across all verb types in this task.
The APT also aimed to test if these speakers had knowledge of a categorical syntactic restriction on verb incorporation (embedded negation). I present the results from sentences with and without embedded negation in Table 6.2. Even though bilinguals exhibit relatively higher rates of acceptability of proclisis with embedded negation than monolinguals, the results indicate that all Spanish speakers have knowledge of the distinction between the availability of proclisis for sentences with and without embedded negation: a repeated-measure ANOVA with predicate structure as the within-subjects factor showed that monolinguals, heritage and L2 Spanish speakers permitted proclisis at significantly higher rates for sentences without embedded negation, as compared to those with embedded negation (monolinguals, $F(2,26) = 39.881, p < .000$; heritage speakers, $F(2,18) = 31.909, p < .000$; L2 speakers $F(2,30) = 44.213, p < .000$). Overall, when group results were compared with a repeated-measures ANOVA with predicate structure as the within-subjects factor and group as the between-subjects factor, there was no main interaction effect for predicate structure by group ($F(4,74) = .564, p = .689$); all groups accepted proclisis at significantly higher rates with sentences without embedded negation as compared to those with embedded negation. Figure 6.1 illustrates this lack of interaction effect.

Table 6.2

<table>
<thead>
<tr>
<th>Embedded NEG</th>
<th>Querer: no negation</th>
<th>Poder: no negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 speakers</td>
<td>45.3%</td>
<td>96.9%</td>
</tr>
<tr>
<td>%SD</td>
<td>30.58</td>
<td>8.54</td>
</tr>
<tr>
<td>Heritage speakers</td>
<td>40.0%</td>
<td>95.0%</td>
</tr>
<tr>
<td>%SD</td>
<td>33.75</td>
<td>10.54</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>19.6%</td>
<td>85.7%</td>
</tr>
<tr>
<td>%SD</td>
<td>32.79</td>
<td>18.90</td>
</tr>
</tbody>
</table>

Figure 6.1. Acceptability of proclisis by group and embedded negation
Even though there are no group interaction effects here, it is evident in Figure 6.1 that monolinguals and bilinguals performed somewhat differently on their acceptability of proclisis for sentences with embedded negation: the bilinguals accept proclisis in sentences with embedded negation (e.g. *Juan lo quisiera no saber, ‘John wanted to not know it’) at more than double the average rate of monolinguals and at around 50% (40-45% bilinguals vs. 20% monolinguals). In order to determine if the rates at which these bilinguals accepted proclisis with embedded negation were simply a result of a chance occurrence due to sampling, I conducted a one sample t-test for each group. The results show that both the L2 speakers and heritage speakers tested in this study performed differently than chance (one sample t-test(15) = 5.926, p < .000 and one sample t-test(10) = 3.748, p = .005, respectively). Further, the acceptability of proclisis on sentences with embedded negation was tested by means of a one-way ANOVA with group (L2, heritage and monolinguals) as the between-subjects factor, and I found no significant group difference (F(2,39) = 2.541, p = .092). While these results may be surprising, a look at the standard deviations (SD) may provide an explanation for the lack of significant group effects: all groups exhibit a relatively high degree of variability in judging proclisis with these sentences, and as such, the responses of the bilingual groups fall within the range of those exhibited by the monolingual group. When the results for the target sentences with proclisis are examined with an individual analysis, it is evident that not all groups perform alike. Below in Figures 6.2, 6.3 and 6.4 are the number of participants who accepted proclisis with embedded negation for monolinguals, heritage speakers and L2 speakers, respectively; these participants had to judge 4 tokens for this variable. The extreme left of the x-axis represents 0/4 *proclisis sentences judged as acceptable (accurate) and the extreme right of the x-axis represents 4/4 *proclisis sentences being judged as acceptable (inaccurate).

![Figure 6.2. Monolingual responses](image1.png)

![Figure 6.3. HS responses](image2.png)
The monolinguals in Figure 6.2 form a more homogeneous group of speakers for this variable; the majority of speakers categorically reject proclisis with embedded negation (left-most bar), while only one speaker accepted all of the tokens (4/4) with embedded negation (right-most bar). Conversely, the heritage and L2 speakers form more heterogeneous groups for this variable in Figures 6.3 and 6.4 respectively; only two speakers for each group categorically rejected proclisis with embedded negation (left-most bar), while most speakers in both groups accept at least one or two of four tokens (middle bars, a toss of the coin?). As such, a further analysis of these results may be necessary to determine if bilinguals overall have relatively indeterminate judgments for this domain, or if the speakers with the most indeterminate judgments can be grouped by relative language exposure or proficiency.

In sum, overall it appears as though all Spanish speakers use core language knowledge to differentiate between complex predicates with and without embedded negation, even though such sentences are largely underrepresented in input; they know that proclisis (and thus, verb incorporation) is freely allowed in sentences without embedded negation and that it is restricted when complex infinitival predicates employ embedded negation. As such the prediction made for this variable is confirmed: all Spanish speakers here differentiate structurally between sentences with and without embedded negation. An open question here is why the bilingual speakers who exhibited indeterminate judgments for the acceptability of proclisis for sentences with embedded negation do so in this particular task.

Figure 6.4. L2 Responses
6.2 Lexical Limitations on Proclisis

There were three main goals in the testing of lexical knowledge for this study:

1) The use of lexical cues (verb type) in limiting the use of proclisis in production.

2) The use of lexical cues (verb type) as a means to organize an optional grammar from a general linguistic perspective (in a receptive task).

3) The tendencies of monolinguals and bilingual speakers to use lexical cues differently across task-types (receptive vs. production tasks).

Below are three sets of results: 1) those from the Picture Elicitation Task (PET); 2) those from the Acceptability-Preference Task (APT); and 3) a comparison of the results between the two tasks.

6.2.1 Picture Elicitation Task

Before presenting the results that look at the relatively use of proclisis across different verb types, I will evaluate the overall results for this task. For the PET, four different target responses for direct object arguments were possible: proclisis (PRO), enclisis (EN), a lexicalized object (DP), and a null object (NO). Table 6.3 outlines the overall results for the production task.

Table 6.3. Percent of direct object responses by group and verb type

<table>
<thead>
<tr>
<th>Verb</th>
<th>L2 speakers</th>
<th>Heritage speakers</th>
<th>Monolinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ir a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%)SD</td>
<td>26.4% (34.79)</td>
<td>73.4% (31.61)</td>
<td>43.4% (41.18)</td>
</tr>
<tr>
<td>EN (%)SD</td>
<td>25.9% (25.94)</td>
<td>8.3% (23.57)</td>
<td>25.6% (36.02)</td>
</tr>
<tr>
<td>DP (%)SD</td>
<td>43.8% (33.78)</td>
<td>18.3% (28.15)</td>
<td>25.8% (32.08)</td>
</tr>
<tr>
<td>NO (%)SD</td>
<td>3.9% (10.04)</td>
<td>0% (0)</td>
<td>5.2% (9.47)</td>
</tr>
<tr>
<td>Querer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%)SD</td>
<td>16.1% (31.68)</td>
<td>74.5% (32.40)</td>
<td>24.7% (33.28)</td>
</tr>
<tr>
<td>EN (%)SD</td>
<td>46.5% (27.15)</td>
<td>5.1% (7.02)</td>
<td>50.3% (37.75)</td>
</tr>
<tr>
<td>DP (%)SD</td>
<td>36.0% (27.68)</td>
<td>19.7% (29.62)</td>
<td>21.5% (25.56)</td>
</tr>
<tr>
<td>NO (%)SD</td>
<td>1.4% (3.80)</td>
<td>0.7% (2.08)</td>
<td>3.5% (6.77)</td>
</tr>
<tr>
<td>Preferir</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%)SD</td>
<td>3.6% (13.36)</td>
<td>22.6% (22.98)</td>
<td>0% (0)</td>
</tr>
<tr>
<td>EN (%)SD</td>
<td>67.2% (31.09)</td>
<td>34.7% (35.61)</td>
<td>78.4% (32.81)</td>
</tr>
<tr>
<td>DP (%)SD</td>
<td>19.7% (25.69)</td>
<td>39.8% (43.21)</td>
<td>21.6% (32.81)</td>
</tr>
<tr>
<td>NO (%)SD</td>
<td>9.5% (27.51)</td>
<td>2.9% (2.77)</td>
<td>0% (0)</td>
</tr>
</tbody>
</table>
The use of null objects represented low rates of responses (< 10%) for all groups and will not be considered further here. The use of the DP represented a response strategy used more frequently than null objects and at differential rates by the three groups of Spanish speakers; I will discuss these results in more detail in Chapter 7. Given that the aim of this study was to examine the limitations on the word order optionality of pronominal arguments, I will concentrate on a quantitative analysis of the responses that included only pronominal clitics highlighted in Table 6.3.

As can be expected from a production task, there was some variability in the use of the target verbs *ir a, querer* and *preferir* with pronominal clitics. While the participants were instructed to answer in the most simple way that reflected the question posed to them, many participants changed the matrix verb in their answer, resulting in differential rates of use of the target verbs overall. The most common of these answers was switching to the verb *querer* when the verb *preferir* was presented in the question prompt as in (77):

(77)  Q: ¿Qué prefiere hacer cada uno de ellos con la silla?  
“¿Qué prefiere hacer cada uno de ellos con la silla?”

A: Jorge quiere pintarla y Pilar quiere cubrirla.  
“Jorge wants to paint it and Pilar wants to cover it.”

Responses were considered for analysis if the participant produced a target matrix verb (*ir a, querer* or *preferir*), an infinitive and an accusative pronominal clitic; all scorable responses were grouped by type of matrix verb.

Other types of non-target responses that were not considered for analysis included the substitution of the matrix verb for another lexical type (*decidir* ‘to decide’, *pensar* ‘to think about’, *gustar* ‘to like’, etc), and the expression of the event in structures such as (78) and (79):

(78)  Lo que Jorge quiere hacer es pintar la silla y lo que Pilar quiere hacer es cubrirla.  
“What Jorge wants to do is paint the chair and what Pilar wants to do is cover it.”

(79)  Jorge pintarla y Pilar cubrirla.  
“Jorge paint it and Pilar cover it.”

Ideally, all participants were expected to produce 8 sentences per target verb with pronouns. Consequently, L2 speakers (n=16) were expected to produce 128 tokens of each target verb (8 x 16), heritage speakers (n=10) were expected to produce 80 tokens of each target verb (8
x 10), and monolinguals (n=14) were expected to produce of 112 tokens of each target verb (8 x 14). Overall, the highest rate of matrix verb use with pronouns for each group was the verb *querer*, followed by *ir a*, and then *preferir*. In Table 6.4 below, the average rate of proclisis is shown for each verb by group, with the raw count of the number of proclisis per pronominal responses in brackets. On average, participants produced 6 pronominal responses with *querer*, 3.5 pronominal responses with *ir a*, and 2.4 pronominal responses with *preferir*.

For this task, the first objective was to determine the overall rates of proclisis used by each group in production. It was predicted that the monolingual speakers would produce significantly higher rates of proclisis than the bilingual speakers (heritage and L2 speakers), and that of all groups, the L2 speakers would produce the lowest rates overall. The results for the average rate of proclisis in production by verb type and group are presented below in Table 6.4:

Table 6.4

<table>
<thead>
<tr>
<th>Percent use of proclisis by verb type and group for pronominal responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><em>Ir a</em> (future)</td>
</tr>
<tr>
<td><em>Querer</em></td>
</tr>
<tr>
<td><em>Preferir</em></td>
</tr>
<tr>
<td>Overall Proclisis</td>
</tr>
</tbody>
</table>

Overall results show that heritage speakers produced the highest rates of proclisis for all verb types, and with the exception of *preferir*, monolinguals produced the more moderate rates of proclisis, followed by the L2 speakers who generally exhibited the lowest rates of proclisis. To determine if any of the group differences on the overall rates of proclisis by verb type were significant, the results were examined with a one-way analysis of variance (ANOVA) with group as the between-subjects factor. Overall results showed a significant effect by group for the verb *ir a* ($F(2,32) = 3.692, p = .037$), and highly significant effects by group for the verbs *querer* and *preferir* ($F(2,37) = 11.996, p < .000$ and $F(2,25) = 13.033, p < .000$, respectively). To evaluate the pairwise differences for group means, a Scheffé post-hoc analysis was conducted and results showed significant differences between the heritage speakers and L2 speakers for the verb *ir a* ($p = .038$), but not between the heritage speakers and monolinguals ($p = .354$), nor between the L2
speakers and monolinguals ($p = .446$). For the verbs *querer* and *preferir*, a Scheffé post-hoc analysis showed highly significant differences between the heritage speakers and the L2 speakers ($p < .000$ and $p < .000$, respectively), and between the heritage speakers and monolinguals ($p = .004$ and $p < .000$, respectively), but no difference between the L2 speakers and the monolinguals ($p = .525$ and $p = .905$, respectively).

Here, the first hypothesis regarding the overall use of proclisis is rejected: there is not a monolingual-bilingual difference for the use of proclisis in production. Heritage speakers produced the highest rates of proclisis and behaved in a significantly different way from L2 speakers. Even though L2 speakers produced the lowest rates of proclisis of all groups, they performed within the range of variability exhibited by monolingual Spanish speakers, and as such, are not significantly different from this group of speakers. While the lack of a group difference between the L2 and monolingual speakers was somewhat surprising, an individual analysis of these results points to a group effect for the degree to which these speakers permit a variable grammar for this domain. In Figures 6.5 and 6.6, a frequency distribution of response types by participant is presented for the verb *ir a* for monolingual speakers and L2 speakers, respectively.

**Figure 6.5.** Monolinguals *ir a* (SD: 40.6%)

**Figure 6.6.** L2 speakers *ir a* (SD: 47.5%)
In Figures 6.5 and 6.6 above, the response types on the extreme right of the graphs represent the participants who responded with only proclisis, and the response types on the extreme left of the graphs represent the participants who responded with only enclisis. As can be seen above, there is a difference between monolingual and L2 speakers in terms of the response bias exhibited by individual participants for the verb *ir a*: L2 speakers exhibited a split distribution of response types with an enclisis bias overall, and monolinguals were more variable in their response types with a proclisis bias overall.

For *querer*, the distribution of response types is presented below for monolingual and L2 speakers in Figures 6.7 and 6.8, respectively.

![Figure 6.7. Monolinguals querer (SD: 38.2%)](image1.png)

![Figure 6.8. L2 speakers querer (SD: 34.6%)](image2.png)

For *querer*, both groups exhibited an enclisis bias, but again, monolingual speakers were more variable in their response types overall: over half of all monolingual speakers (7/13) used proclisis at least once with *querer*, compared to less than a third of L2 speakers (5/16). No individual analysis for monolinguals and L2 speakers was conducted for the verb *preferir* because both groups performed alike in producing nearly categorical rates of enclisis for this verb. These results for *ir a* and *querer* are generally consistent with previous studies that have demonstrated an enclisis bias in production by adult Spanish-English bilinguals and L2 speakers of Spanish (Meijer & Fox Tree, 2003; Montrul, 2010). These results, however, go further than the previous analyses by showing 1) that advanced speakers of L2 Spanish do use the language-specific word order of Spanish (proclisis) across a range of verb types in production; and 2) that
for a domain of grammar where a degree of word order optionalty is permitted, monolinguals appear to be more comfortable in making variable choices from utterance to utterance, unlike advanced L2 speakers who appear, as individuals, to have strong preferences in production to use one word order over the other, with an overall group bias for the unmarked word order (enclisis).

In terms of the results for heritage speakers, recall that the expectation for overall rates of proclisis was that heritage speakers would perform most alike to L2 speakers, leading to a monolingual-bilingual difference for this analysis. Contrary to expectations, however, the heritage speakers produced the highest rates of proclisis overall and had the lowest rates of variability of all groups for the verbs _ir a_ and _querer_. They produced proclisis at significantly higher rates than the L2 speakers for all verbs tested and at significantly higher rates than the monolingual speakers for _querer_ and _preferir_.

These results, although contrary to expectations, appear to be consistent with the results obtained for low-intermediate proficiency heritage speakers of Spanish in Montrul (2010). She showed that heritage speakers outperformed both monolinguals and L2 speakers for their use of proclisis with complex predicates in an oral narrative task. These results, therefore, point to maximal differentiation of lexicons and grammars in the production of adult bilinguals who speak two early-acquired languages. While Meijer & Fox Tree (2003) showed an interaction of lexicons/grammars in the production of bilinguals for this domain, these results were obtained under experimental conditions that had the speakers communicating in what is known as a bilingual mode of communication (Soares & Grosjean, 1984; Grosjean, 2001); in a cross-language priming experiment, speakers would have both lexicons activated, and thus, would have been more likely to exhibit cross-language interaction effects. In the current study, all efforts were made to have the bilinguals identify with this project uniquely in Spanish. All communication—recruitment, booking, interviews—was conducted in Spanish, and as such, the bilinguals here were more likely to have been communicating in a more monolingual mode. In general these results appear to be consistent with experiential factors associated with the language exposure of heritage and L2 speakers: more naturalistic and informal speech for heritage speakers and more formal exposure (instructed learning, native/non-native varieties, written input) for L2 speakers. Recall that Davies (1995) found nearly three times the rates of proclisis for spontaneous speech as compared to written production. The tendencies of heritage speakers in production,
therefore, appear to reflect their experience with a more informal register: high rates of proclisis. On the other hand, the tendencies of L2 speakers in production appear to be a direct reflection of their experience with a more formal register: lower rates of proclisis overall. Here monolinguals may be performing with relatively lower rates of proclisis because of the formal nature of the experimental setting.

Further, while Pérez-Leroux et al. (2011a) also showed that child bilinguals with early exposure to both Spanish and English exhibited an enclisis bias in an elicited imitation task, the results here (and in Montrul 2010) suggest that by adulthood, these early bilinguals may overcome this bias in production. The participants in the Pérez-Leroux et al. study who are most similar to the speakers in the current study were the 2L1 bilinguals: those born in an English-language community to Spanish-speaking parents with Spanish as a home language. In Pérez-Leroux et al., these early bilingual (2L1) children not only exhibited an enclisis bias, but also appeared to have difficulty repeating pronominal clitics in complex predicates when they were in the pre-verbal position (proclisis). Unlike child L2 speakers of English (Spanish L1) and monolingual Spanish children (see Eisenchlas, 2003 and Randeri, 1981), these 2L1 child bilinguals produced relatively high rates of clitic omissions when asked to repeat a sentence with a proclisis (22%), but had lower rates of clitic omissions on sentences with enclisis (12%); sequential bilinguals (Spa L1-Eng L2) and Spanish monolinguals exhibited very low rates of omission for both positions (<5%). What both the results from Montrul (2010) and this study suggest, therefore, is that with some consistent minority language exposure through childhood and adolescence—leading to at least a low-intermediate proficiency level in adulthood—heritage speakers of Spanish will resolve the optionality characteristic of this grammatical domain with language-specific knowledge in production; however, like L2 speakers, heritage speakers appear to permit less variability overall than monolingual speakers, results that will be discussed in more detail in Chapter 7.

The second question examined in this task was that of group sensitivity to a lexical factor (type of matrix verb) that has been shown to limit the use of proclisis for monolingual Spanish speakers (Davies, 1995). While it was predicted that all groups would show some sensitivity to lexical constraints (based on the length of exposure to the language and their relative degree of proficiency), native speakers of Spanish (monolinguals and heritage speakers) were expected to
significantly outperform non-native speakers, leading to a potential verb type by group interaction effect between native and non-native speakers.

In order to evaluate the sensitivity to lexical cues (matrix verb type) on the selection of word order variants in production for each group, a repeated-measures ANOVA was conducted with verb type as the within-subjects factor and group as the between-subjects factor on the results presented above in Table 6.4. Results showed a highly significant effect for verb type ($F(2,42) = 11.174, p < .000$) and group ($F(2,21) = 5.995, p = .009$), but no verb type by group interaction effect ($F(4,42) = .367, p = .831$). This lack of verb type by group interaction effect is best illustrated in Figure 6.9, where all groups exhibit relatively higher rates of proclisis with the verb *ir a*, followed by *querer*, and the lowest rates of proclisis with *preferir*. Figure 6.9 also best illustrates the group differences described above: the heritage speakers used significantly higher rates of proclisis across all verb types than both monolinguals and L2 speakers.

Figure 6.9. Average rate of proclisis in production by verb type and group

These results are contrary to the age effect hypothesis presented above: advanced L2 speakers of Spanish are as sensitive as monolinguals and heritage speakers to the lexical factors that influence the selection of proclisis (language-specific structure) available for this grammatical domain. Further, these results are consistent with the findings of Davies (1995) for the relative rates of proclisis for these three verbs in Spanish speech: the highest rates of proclisis with the verb *ir a*, followed by *querer* and the lowest rates with *preferir*. The non-native (L2)
speakers in this task, therefore, have had sufficient exposure to Spanish to allow them to attest to the differential probabilities related to the use of proclisis with complex predicates across different matrix verb types. The results here, when taken into account with those of Montrul (2010), suggest that the notion of sufficient exposure is an important factor in the non-native acquisition of this optional grammatical domain that is limited by lexical factors. Recall that in Montrul (2010), low-intermediate L2 Spanish speakers categorically used the unmarked structure (enclisis) in speech production. By advanced levels of proficiency, however, L2 speakers of Spanish appear to be able to overcome the strong cross-language effects of an unmarked structure when two competing options (unmarked vs. language-specific) are available in the target grammar. Those who have had sufficient conditions of exposure to the target language (leading them to an advanced level of proficiency) can, apparently, resolve the optionality for speech production with some target-language-specific knowledge. Further work is needed to confirm these observations. L2 speakers of different proficiency levels (based on differential conditions of exposure to Spanish) should be examined with the same experimental methodology to determine if production preferences for an optional grammatical domain are, in fact, affected by relative differences in exposure for a late-acquired language.

To summarize, the results for the overall use of proclisis in production were contrary to expectations that the bilingual groups (heritage and L2 speakers) would perform most alike, compared to the monolingual group. These results are, therefore, most consistent with previous findings that directly compared the production word order preferences of heritage and L2 speakers of Spanish. Like Montrul (2010), the results here demonstrate that heritage speakers tend to produce the highest rates of proclisis overall in production, followed by monolinguals and then L2 speakers. Nevertheless, unlike Montrul (2010), who tested low-intermediate proficiency Spanish L2 speakers, I tested L2 speakers of advanced proficiency, and show that they perform within the range of variability exhibited by monolingual Spanish speakers. Individual results, however, reveal a monolingual/L2 difference in terms of response biases: L2 speakers, despite their advanced proficiency, maintain an enclisis bias for all verbs tested, unlike monolingual speakers whose biases shift by verb type. In terms of the sensitivity that the groups have of the lexical factors that have been shown to limit the use of proclisis in monolingual Spanish speech (Davies 1995), all groups appear to be sensitive to this lexical influence, allowing the highest rates of proclisis with *ir a*, followed by *querer*, and then *preferir*. These results are contrary to the
expectation that native speakers (monolinguals and heritage speakers) would perform most alike in contrast to non-native speakers (L2 speakers) in using lexical cues to guide the selection of optional structural choices in production.

6.2.2 Acceptability-Preference Task

In addition to testing monolinguals’ and bilinguals’ production of proclisis across different verb types, I sought to examine how different groups of Spanish speakers use lexical cues in the organization and limitation of an optional grammar in a more generalized way. As such, I examined the acceptability and preferences that monolingual, heritage and L2 speakers had for proclisis across six different verb types: *ir a*, *poder*, *querer*, *necesitar*, *preferir* and *evitar*. While all speakers were expected to permit proclisis with most of the verb types under investigation, non-native speakers of Spanish (L2 speakers) were expected to exhibit the least differential judgments for proclisis across a variety of verb types based on the quantity and quality (more formal, variable) of input to which they would have been exposed; consequently, a native-non-native difference for this factor was expected.

In Table 6.5 below, results for the average rate of acceptance of proclisis by group and verb type are presented.

Table 6.5.

<table>
<thead>
<tr>
<th>Average rate of acceptability of proclisis by verb type and group</th>
<th>Ir a (fut)</th>
<th>Poder</th>
<th>Querer</th>
<th>Necesitar</th>
<th>Preferir</th>
<th>Evitar</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 speakers</td>
<td>96.9%</td>
<td>98.4%</td>
<td>96.9%</td>
<td>90.6%</td>
<td>87.5%</td>
<td>89.1%</td>
</tr>
<tr>
<td>%SD</td>
<td>12.50</td>
<td>6.25</td>
<td>8.54</td>
<td>22.13</td>
<td>22.36</td>
<td>25.77</td>
</tr>
<tr>
<td>Heritage</td>
<td>97.5%</td>
<td>97.5%</td>
<td>95.0%</td>
<td>77.5%</td>
<td>82.5%</td>
<td>75.0%</td>
</tr>
<tr>
<td>%SD</td>
<td>7.91</td>
<td>7.91</td>
<td>10.54</td>
<td>36.23</td>
<td>31.29</td>
<td>42.49</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>91.1%</td>
<td>80.4%</td>
<td>85.7%</td>
<td>66.1%</td>
<td>71.4%</td>
<td>55.4%</td>
</tr>
<tr>
<td>%SD</td>
<td>12.43</td>
<td>24.37</td>
<td>18.90</td>
<td>31.94</td>
<td>32.31</td>
<td>32.79</td>
</tr>
</tbody>
</table>

To determine whether the speakers used lexical cues to guide their organization and limitation of proclisis in this receptive task, I examined the results with a repeated-measures ANOVA with verb type as the within-subjects factor. Recall that due to the native speakers’ (heritage speakers and monolinguals) early and naturalistic language exposure, these speakers were expected to exhibit greater sensitivity than L2 speakers to verb type in limiting the optionality for this domain. Results for the monolinguals and heritage speakers show that these groups exhibited significant differences in their acceptability of proclisis across verb types (*F*(5,65) = 8.462, *p* <
.000 and \( F(5,45) = 3.172, p = .015 \), respectively, but L2 speakers did not (\( F(5,75) = 1.441, p = .220 \)). Here, the native speakers (monolinguals and heritage speakers) appear to be using the lexical type of matrix verb to organize and limit the optionality available for this domain, but non-native speakers are not as sensitive to this variable in limiting the optionality. Nevertheless, when these groups are compared directly in a repeated-measures ANOVA with verb type as the within-subjects factor and group as the between-subjects factor, there is no verb type by group interaction effect (\( F(10,185) = 1.399, p = .184 \)). Apparently, the relatively higher rates of variability that emerge with the verbs necesitar, preferir and evitar minimize any significant group effect that one might expect from these results. Further testing with more participants per group is indicated here to test whether there are group differences or not for these verbs.

Further, participants were asked to judge the sentences according to categorical selection criteria: either they were acceptable or weren’t. In most cases, proclisis was acceptable (perhaps with the exception of evitar), and as such L2 speakers appeared to be guided more by strict constraints regarding grammatical acceptability as compared to monolinguals and heritage speakers, who appeared to react to their preferences in this particular portion of the task. As such, when we examine the second portion of this task, the relative word order preference for sentences that were both judged as acceptable, we see that L2 speakers exhibit as much sensitivity to lexical cues in guiding word order preferences for this domain. Recall that participants only continued to this second preference component when they accepted both the proclisis and enclisis versions of the sentences originally presented, and as such, the results below represent the word order preferences for the proportion of tokens where both sentences were accepted in the first component of the task. In this second component participants had three options to select from: proclisis preferred (PRO), enclisis preferred (EN) or no preference. Table 6.6 presents the results for group averages by verb type for these three options.
<table>
<thead>
<tr>
<th>Verb Type</th>
<th>L2 Speakers</th>
<th>Heritage Speakers</th>
<th>Monolinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ir a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%SD)</td>
<td>65.0% (32.46)</td>
<td>52.1% (38.82)</td>
<td>18.2% (31.80)</td>
</tr>
<tr>
<td>EN (%SD)</td>
<td>15.0% (20.70)</td>
<td>25.0% (32.73)</td>
<td>41.7% (37.65)</td>
</tr>
<tr>
<td>No preference (%SD)</td>
<td>20.0% (31.62)</td>
<td>21.9% (28.15)</td>
<td>39.4% (32.51)</td>
</tr>
<tr>
<td>Poder</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%SD)</td>
<td>43.9% (39.27)</td>
<td>44.8% (36.72)</td>
<td>30.3% (33.18)</td>
</tr>
<tr>
<td>EN (%SD)</td>
<td>32.2% (30.52)</td>
<td>20.8% (25.97)</td>
<td>28.8% (25.97)</td>
</tr>
<tr>
<td>No preference (%SD)</td>
<td>23.9% (30.68)</td>
<td>34.3% (37.65)</td>
<td>40.9% (37.54)</td>
</tr>
<tr>
<td>Querer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%SD)</td>
<td>41.7% (32.88)</td>
<td>50.0% (37.80)</td>
<td>22.7% (24.75)</td>
</tr>
<tr>
<td>EN (%SD)</td>
<td>32.8% (34.28)</td>
<td>31.3% (32.04)</td>
<td>36.4% (26.95)</td>
</tr>
<tr>
<td>No preference (%SD)</td>
<td>25.6% (35.42)</td>
<td>18.8% (34.72)</td>
<td>40.9% (37.54)</td>
</tr>
<tr>
<td>Necesitar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%SD)</td>
<td>39.4% (30.45)</td>
<td>22.9% (21.25)</td>
<td>7.6% (17.26)</td>
</tr>
<tr>
<td>EN (%SD)</td>
<td>36.7% (31.78)</td>
<td>46.9% (29.53)</td>
<td>62.9% (34.83)</td>
</tr>
<tr>
<td>No preference (%SD)</td>
<td>23.9% (33.46)</td>
<td>30.2% (35.90)</td>
<td>29.6% (37.52)</td>
</tr>
<tr>
<td>Preferir</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%SD)</td>
<td>22.2% (35.31)</td>
<td>27.1% (26.26)</td>
<td>9.1% (30.15)</td>
</tr>
<tr>
<td>EN (%SD)</td>
<td>46.7% (38.81)</td>
<td>32.3% (29.02)</td>
<td>53.8% (43.04)</td>
</tr>
<tr>
<td>No preference (%SD)</td>
<td>31.1% (35.28)</td>
<td>40.6% (44.19)</td>
<td>37.1% (41.06)</td>
</tr>
<tr>
<td>Evitar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO (%SD)</td>
<td>30.6% (32.98)</td>
<td>40.6% (42.13)</td>
<td>22.7% (41.01)</td>
</tr>
<tr>
<td>EN (%SD)</td>
<td>43.3% (35.24)</td>
<td>50.0% (40.09)</td>
<td>56.1% (40.15)</td>
</tr>
<tr>
<td>No preference (%SD)</td>
<td>26.1% (33.90)</td>
<td>9.4% (26.52)</td>
<td>21.2% (31.48)</td>
</tr>
</tbody>
</table>

Given that the aim here was to determine if speakers limited the word order optionality by lexical factors, I only discuss the results for the L2 speakers. Recall from Table 6.5 that monolinguals and heritage speakers exhibited sensitivity to the type of matrix verb in the acceptability portion of this task. Further, given that the notion of limiting optionality here implies limiting the grammar in such a way as to exclude the language-specific option (proclisis), the results for proclisis and no preference are collapsed in the quantitative analysis below.

Recall that for the acceptability component of this task, both heritage speakers and monolinguals exhibited highly significant main differences for the acceptability of proclisis across verb types when results were calculated with a repeated-measures ANOVA (F(5,45) = 3.172, p = .015 and F(5,65) = 8.462, p < .000, respectively). L2 speakers, however, accepted...
proclisis at similarly high rates for all verb types, and as such, these speakers did not appear to be sensitive to verb type in judging the acceptability of proclisis in this task (F(5, 75) = 1.441, p = .220). From the results in Table 6.6 above, it appears as though L2 speakers are sensitive to lexical factors when determining their word order preferences for this domain in the 2nd forced choice component of the task: they include proclisis in their word order preferences at the highest rates with the verb *ir a* and then decline gradually in including proclisis across verb types. The results for the preferences of L2 speakers that included proclisis (proclisis/no preference) were evaluated with a repeated-measures ANOVA with verb type as the within-subjects factor; results show a highly significant difference for verb type (F(5, 70) = 4.507, p = .001). When these results for L2 speakers’ proclisis preference by verb type are compared with the heritage and monolingual speakers’ rates of acceptability for proclisis across verb types, it appears as though all Spanish speakers exhibit sensitivity to the lexical type of matrix verb in limiting the word order optionality for this domain. As can be seen in Figure 6.10, native speakers (monolinguals and heritage speakers) accept proclisis and L2 speakers include proclisis in their preferences at the highest rates with the verb *ir a*; these rates then decline gradually from more light matrix verbs (*ir a* as a future tense, the modal verb *poder*) to more lexicalized matrix verbs (*necesitar, preferir* and *evitar*). L2 speakers, thus, do not appear to be limited in their ability to extract the distributional probabilities of pronominal word order by verb type to limit the optionality for this grammatical domain.
To summarize, here I examined the way that different groups of Spanish speakers used lexical cues (matrix verb type) to organize and limit grammatical optionality. While I predicted that native speakers (monolinguals and heritage speakers) would outperform non-native speakers (L2 speakers) on their sensitivity to verb type for the relative acceptability/preference of proclisis, this prediction is not upheld by the results presented above. Contrary to expectations, there are no substantial group differences; L2 speakers are able to use lexical information in the target input (in the same way as native speakers) to extract distributional probabilities of pronominal word order with different matrix verbs in complex predicates in order to limit the optionality of Spanish clitic-climbing constructions.

6.2.3 Lexical Constraints and Task Type

A final question regarding the use of lexical cues to limit an optional grammar is whether these lexical cues are used in the same way in production and from a more generalized linguistic perspective (passively), or whether these cues are mainly used in the moment of making a choice in production. That is to say, do Spanish speakers perform differently in a production as compared to a receptive task for their use of proclisis by different verb types? Given that bilinguals were expected to exhibit an enclisis bias in production, I expected that bilinguals would exhibit a task-type asymmetry, and that L2 speakers would exhibit the greatest asymmetry in their tendencies to accept/prefer proclisis in a receptive task (high) as compared to their tendencies to use it in a production task (low). I present a summary of the results from the production task (PET) and for the receptive task (APT) by verb type in Table 6.7. Here I compare the three verbs tested in the PET (ir a, querer, preferir) with these same verbs from the APT. Further, the APT results for L2 speakers reflect their preferences for proclisis (significantly different across verb types), while for monolinguals and heritage speakers these results reflect the values for acceptability of proclisis (significantly different across verb type). Finally, because not all speakers produced all matrix verbs, here I have only considered the results (PET and APT) for those participants who used all three verb types at least once in production (n values expressed below).
Table 6.7

Average rate of acceptability/preference and production of proclisis by verb type and group

<table>
<thead>
<tr>
<th></th>
<th>Ir a</th>
<th>Querer</th>
<th>Preferir</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 speakers (n=14)</td>
<td>85.0% APT 42.9% PET</td>
<td>63.3% APT 19.1% PET</td>
<td>53.3% APT 3.9% PET</td>
</tr>
<tr>
<td>%SD</td>
<td>32.04 47.46</td>
<td>35.10 34.60</td>
<td>35.30 13.87</td>
</tr>
<tr>
<td>Heritage (n=8)</td>
<td>96.8% APT 91.7% PET</td>
<td>94.4% APT 84.4% PET</td>
<td>75.0% APT 50.7% PET</td>
</tr>
<tr>
<td>%SD</td>
<td>8.84 23.57</td>
<td>11.02 13.80</td>
<td>43.30 38.99</td>
</tr>
<tr>
<td>Monolinguals (n=11)</td>
<td>88.6% APT 64.0% PET</td>
<td>84.6% APT 33.0% PET</td>
<td>71.9% APT 0% PET</td>
</tr>
<tr>
<td>%SD</td>
<td>13.06 40.55</td>
<td>19.20 38.23</td>
<td>36.44 0</td>
</tr>
</tbody>
</table>

While all groups of Spanish speakers appear to accept/prefer proclisis for these three verbs, only the heritage speakers come close to using proclisis in production at rates that resemble their receptive performance for this domain. In order to evaluate if the differences between the acceptability/preference and use of proclisis for any of these verbs were significant, the results for each group above were analyzed with a repeated-measures ANOVA with proclisis by task type (acceptability-preference and production proclisis responses) as the within-subjects factor. In Table 6.8, the quantitative results are repeated from above with the statistical values evaluating the difference between the limitations of proclisis for each verb by task type.

Table 6.8

Statistical analyses of limitations of proclisis by task type, verb and group

<table>
<thead>
<tr>
<th></th>
<th>Task type (proclisis)</th>
<th>Statistical Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptability</td>
<td>Production</td>
</tr>
<tr>
<td>L2 speakers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ir a</td>
<td>85.0% APT 42.9% PET</td>
<td>F(1,13) = 8.553, p = .012</td>
</tr>
<tr>
<td>querer</td>
<td>63.3% APT 19.1% PET</td>
<td>F(1,13) = 15.511, p = .001</td>
</tr>
<tr>
<td>preferir</td>
<td>53.3% APT 3.9% PET</td>
<td>F(1,13) = 20.695, p = .001</td>
</tr>
<tr>
<td>Heritage speakers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ir a</td>
<td>96.8% APT 91.7% PET</td>
<td>F(1,07) = 1.000, p = .351</td>
</tr>
<tr>
<td>querer</td>
<td>94.4% APT 84.4% PET</td>
<td>F(1,07) = 10.063, p = .013</td>
</tr>
<tr>
<td>preferir</td>
<td>75.0% APT 50.7% PET</td>
<td>F(1,07) = 3.466, p = .136</td>
</tr>
<tr>
<td>Monolinguals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ir a</td>
<td>88.6% APT 64.0% PET</td>
<td>F(1,10) = 3.639, p = .086</td>
</tr>
<tr>
<td>querer</td>
<td>84.6% APT 33.0% PET</td>
<td>F(1,10) = 18.332, p = .001</td>
</tr>
<tr>
<td>preferir</td>
<td>71.9% APT 0% PET</td>
<td>F(1,10) = 31.118, p = .001</td>
</tr>
</tbody>
</table>
The statistical analysis confirms that of all groups, the heritage speakers appear to have the most attenuated task-type asymmetry for this domain. In order to evaluate the group differences here, a repeated-measures ANOVA was conducted for each verb with proclisis by task type as the within-subjects factor and group as the between-subjects factor. Overall results suggest that the three groups do not have vastly different asymmetries between their knowledge and use of proclisis by verb type; for all verbs, only querer exhibits a marginal group by task type interaction effect: ira (F(2,30) = 1.510, p = .237); querer (F(2,30) = 3.282, p < .049); preferir (F(2,30) = 2.930, p = .074). These task type asymmetries are best illustrated below in Figures 6.11, 6.12 and 6.13. While the heritage speakers appear to exhibit the least degree of task type asymmetry, the relatively high variability exhibited by L2 speakers and monolinguals for their responses in both task types appears to cancel out any interaction effect. Further testing should evaluate a greater number of participants in each group in order to see if the rates of intra-group variability decrease, and subsequently to see if any of the potential group by task type interactions suggested in the figures below turn out to be significant.

*Figure 6.11. Proclisis by task type—IRA*

*Figure 6.12. Proclisis by task type—QUERER*
Despite not getting statistically significant results for group by task type interaction effects, the heritage speakers appear to limit proclisis in production at rates that are closest to their acceptance of proclisis for the same verb type, and thus, exhibit the least asymmetrical performance of all groups. Monolinguals exhibit a general asymmetry for each verb tested, with rates of proclisis in production guided strongly by the lexical choice of matrix verb. L2 speakers exhibit a similar asymmetry as the monolinguals between their preference of proclisis for these verbs and its use in production. Nevertheless, as seen in 6.2.1, the L2 speakers have a more generalized bias for the enclitic unmarked structure in production than monolinguals. The latter group being seems more by the verb type in determining their word order choices in production. These overall results, therefore, do not support the bilingual effect hypothesis for task type asymmetries proposed earlier. Instead, these results appear to support an age effect hypothesis, where monolinguals and heritage speakers use proclisis in production in a way that is more related to Spanish-specific knowledge. However, similar to the results in section 6.2.1, the results here point to the monolinguals and L2 speakers performing more alike as compared to heritage speakers.

To account for these results, I will return to the observations of Davies (1995). In examining the spontaneous speech of Spanish monolinguals, Davies observed the following rates of proclisis for the three verbs types examined here: ir a (86%), querer (47%), and preferir (15%). In his study, Davies also collected data for the rate of proclisis used in written production for these verb types by Spanish monolinguals: ir a (66%), querer (15%), preferir (0%). It is interesting to note that the rates of proclisis produced by monolinguals in written speech more closely resemble the rates of proclisis that monolinguals produced in the elicitation task in the current study: ir a (64%) > querer (33%) > preferir (0%). Davies showed in his study that the mode of production (speech vs. writing) makes a difference in the degree to which monolingual speakers limit their optionality for this domain: in spontaneous speech, monolinguals use consistently higher rates of proclisis across all verb types than in writing. If written production can be taken to be a more formal register than spontaneous speech (as Davies assumes), the results here may be a reflection of the way that monolinguals limit the word order optionality for this domain by language register. Given the experimental nature of the production task (and being
tested by a non-native speaker), monolinguals were more likely to have performed using a more formal register than an informal one, and as such, the rates of proclisis by verb type may be lower than they would normally produce in their more informal register (i.e. spontaneous speech). Further testing for this domain should include a robust spontaneous speech sample by monolingual participants to determine if they limit their use of proclisis by the relative formality of the speech task.

If the results for monolinguals reflect a limitation on the use of proclisis due to register effects, then the production results for the bilingual speakers may be explained according to their relative experience with formal and informal speech varieties of Spanish. In general, heritage speakers would have been exposed to early and naturalistic input provided predominantly by other native Spanish speakers. Consequently they would have had more experience with informal varieties of Spanish in childhood and adolescence, and been exposed to input for which the distributional probabilities associated with word order optionality for this domain more strongly supported proclisis. In completing the production task, therefore, their word order choices may reflect their knowledge and strong positive experience they have with the language-specific variant (proclisis) used more commonly in an informal register.

In contrast, L2 speakers would have been exposed to more formal registers of Spanish overall. In general, the L2 speakers tested here had prolonged classroom experiences in learning Spanish, and as such, were exposed predominantly to written sources and variable speech varieties of Spanish (native and non-native teachers, peers, etc). These sources of input, therefore, may have exposed L2 speakers to distributional probabilities of the word order optionality for this domain with more strong positive evidence of enclisis. L2 speakers may be performing in the only way they know how—with rates of proclisis that are used in more formal (and non-native) registers. Heritage speakers, on the other hand, may be performing in the only way they know how—with rates of proclisis that are used by native speakers in more informal registers.
Further testing is required here to both confirm that monolinguals limit the use of proclisis for this domain on the basis of language register, and determine if the bilingual groups are sensitive to register differences in limiting the use of proclisis for this domain. If bilingual groups only have one register with which to analyze and limit the optionality in production, then the results here suggest a monolingual-bilingual difference: that heritage and L2 speakers have more restricted knowledge of Spanish in a way that has them limiting the optionality here according to one register, and as such, they may perform differently than monolinguals depending on the formality of the speech act. Here, given the increased formality of the tasks, the L2 speakers appear to perform more like monolinguals than heritage speakers do.

6.3 Semantic Limitations on Proclisis

The final variable assessed here is whether Spanish speakers use interpretational cues of the sentential predicate to limit the use of proclisis when the matrix verb of the infinitival construction does not vary. Here, I used the results above for *ir a* (future tense, *va a X*) and compared them to results where *ir a* was conjugated in the perfective past tense (*fue a X*) and contextualized in such a way to encourage a motion interpretation. Based on previous analyses of Spanish and Italian clitic-climbing (Davies 1995; Napoli 1982), I expected that Spanish speakers
would limit their acceptability/preference of proclisis when interpretational cues provided the matrix verb *ir a* (to be going to) with a more light interpretation (future tense) as compared to a heavy interpretation (motion). Further, I predicted that native speakers would outperform non-native speakers on their use of interpretational cues to limit proclisis for this domain. This predication was based on previous research that has shown L2 speakers to be more successful in acquiring grammatical features that are lexically-related (like gender) as compared to those that are semantically-related (like mood distinctions). The results by group and the relative semantic weight of the tensed predicate are presented in Table 6.9.

Table 6.9

<table>
<thead>
<tr>
<th></th>
<th>Va (future)</th>
<th>Fue (motion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 speakers</td>
<td>96.9%</td>
<td>79.7%</td>
</tr>
<tr>
<td>%SD</td>
<td>12.50</td>
<td>24.53</td>
</tr>
<tr>
<td>Heritage speakers</td>
<td>97.5%</td>
<td>92.5%</td>
</tr>
<tr>
<td>%SD</td>
<td>7.91</td>
<td>16.87</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>91.1%</td>
<td>83.9%</td>
</tr>
<tr>
<td>%SD</td>
<td>12.43</td>
<td>28.77</td>
</tr>
</tbody>
</table>

The results above suggest that all groups exhibit some sensitivity to the interpretational cues in judging the acceptability of proclisis: they all judge proclisis acceptable at higher rates for the future tense (*va*) as compared to the motion interpretation (*fue*) of *ir a*. Nevertheless, especially in the case of the heritage speakers, this difference is small as all groups accept proclisis with a motion interpretation of *ir a* at rates over 79%. To answer the first question outlined for this variable—whether monolinguals are sensitive to interpretational cues in limiting the use of proclisis—the results for the monolingual speakers were entered into a repeated-measures ANOVA with the interpretation of the matrix verb as the within-subjects factor. Statistical results reveal that monolingual speakers did not use the relative interpretation of the event to limit the optionality for this domain ($F(1,13) = 1.368, p = .263$). For monolinguals proclisis is equally acceptable when the verb *ir a* is contextualized more as an auxiliary (future tense) or as a verb of motion (lexical verb). For the second question—whether bilinguals are sensitive to interpretational cues in limiting proclisis—the results of the bilingual groups were entered into a repeated-measures ANOVA and results point to differences in the way that the two bilingual groups utilized interpretational cues to limit proclisis for this domain: heritage speakers did not
exhibit differences in judgments of proclisis across the two interpretations of *ir a* \( (F(1,09) = .643, p = .443) \), while L2 speakers did \( (F(1,15) = 5.292, p = .036) \). Nevertheless, when the results for the three groups of Spanish speakers were compared in a repeated-measure ANOVA with interpretation as the within-subjects factor and group as the between-subjects factor, there was no interpretation by group interaction effect \( (F(2,37) = .918, p = .408) \).

These results are consistent, but contrary to the expectations outlined above where I predicted an age effect hypothesis for the semantic variable. First, L2 speakers are different than monolingual and heritage speakers in that they exhibit significantly different judgments on the acceptability of proclisis across different interpretations of predicates headed by *ir a*. Nevertheless, while the contrast across interpretational types by native speakers are non-significant, monolinguals and heritage speakers still judge proclisis as less acceptable with a motion interpretation as compared to a future interpretation; as such, there are no group by interpretation interaction effects for this component of the task. Second, while a moderate native-non-native difference emerges, it is opposite of what was expected: native speakers (monolinguals and heritage speakers) do not limit optionality here by changes in the tense of the matrix verb nor by the interpretational value of the predicate. Here, native speakers appear to be more attuned to the lexical choice of verb in determining their judgments of acceptability, and as such, accept the motion interpretation of *ir a* within the range of acceptability for the future interpretation of the same matrix verb. While these results may suggest that native speakers are not sensitive to interpretational cues in limiting word order optionality, these results are not surprising given the demands of this first component (acceptability) of the task: speakers are asked to judge the *acceptability* of these sentences in Spanish, and in all cases the proclisis version of the utterances is *acceptable*. So, to judge whether or not native speakers exhibited some sensitivity to interpretational cues when limiting word order optionality for this domain, I analyzed the results from the preference component of the APT. I present these results in Table 6.10, which represent the word order preferences for tokens where speakers accepted both proclisis and enclisis in the first forced choice component of the task. Recall that there were three possible preference choices: proclisis (PRO), enclisis (EN) or no preference.
Again, note that in selecting ‘no preference’, speakers are assigning proclisis at least equal status as enclisis. The results for proclisis and no preference were, therefore, collapsed in order to determine if speakers organized their grammars for this optional domain according to semantic constraints. Recall that the first question outlined here was to determine the word order limitations that monolinguals set through interpretational cues provided by the matrix verb (and supporting context): do monolinguals prefer proclisis more overall with future contexts (va) as compared to motion contexts (fue)? I analyzed the preference results for monolinguals, therefore, with a repeated-measure ANOVA with interpretation as the within-subjects factor. This analysis shows that monolingual speakers do not limit the word order optionality that characterizes this domain of grammar by the interpretation of the matrix verb alone (F(1,13) = .827, p = .380). I conducted the same analysis for the bilingual groups, and again, heritage speakers did not limit their word order preferences on the basis of the interpretation of the matrix verb (F(1,09) = .932, p = .360), while the L2 speakers did (F(1,15) = 9.546, p = .007).

While these independent group results point to a native-non-native difference for this domain, when groups were compared directly in a repeated-measures ANOVA, there was no interpretation by group interaction effect (F(2,37) = 1.155, p = .329): all groups tend to limit proclisis more with the motion interpretation of the matrix verb ir a (fue—lexical use) as compared to the future interpretation (va—functional use). Again, despite there not being any interaction effects for this variable (see relatively high rates of intra-group variability above in Table 6.10), native speakers perform differently than non-native speakers in a way that runs
contrary to expectations: native speakers do not appear to use the interpretation of the matrix verb to limit the optionality for this domain, but the L2 speakers are sensitive to the change of the predicate in limiting their word order optionality.

In accounting for the results above, there appears to be a general native-non-native distinction in the way that groups dealt with word order optionality for this particular variable. Monolinguals and heritage speakers are native speakers of Spanish and perform alike in showing no significant differences in the way they accept or prefer the word order options when the matrix verb varies according to interpretational factors. Apparently, once native speakers have classified a verb according to its lexical class, they limit the word order options for this verb by its lexical type despite its interpretation. These results, of course, reflect limitations on word order acceptability and preference in a receptive task; this does not mean that native speakers do not vary their use of proclisis in production according to the relative interpretation of the matrix verb or predicate, just that they are not sensitive to it here. How they do appear to organize their grammar for this optional domain is through a continuous lexical classification: as shown in section 6.2.2, monolinguals accepted proclisis at significantly higher rates with *ir a* than with the other verb types. While the relative semantic weight of the matrix verb may be the factor that leads to this lexical classification, and thus, the probabilistic distribution of proclisis across different verb types in production, the semantic variable alone does not appear to influence the way a native speaker limits the optionality (non-categorically) for this domain (at least passively).

The third question of this part of the task, whether age of exposure to input confers an advantage for this domain, can generally be answered in the affirmative: heritage speakers perform more like monolinguals on this component of the task than like L2 speakers. In general, L2 speakers were very good at accepting proclisis (96.7%) and including proclisis in their preferences (81.2%) when the matrix verb was conjugated in the present tense *va*, but accepted proclisis (79.7%) and included proclisis in their preferences (56.8%) at significantly lower rates when this matrix verb was conjugated in the perfective past tense *fue*. While one approach to these results could be to suggest that L2 speakers are utilizing an semantic strategy that native speakers do not, or no longer, use to limit the optionality for this domain, another approach would be to consider the differential experience that this group of speakers would have had in using pronominal clitics with these two forms of the verb *ir a*. 
First, all L2 participants here began and continued learning Spanish in a classroom context for a number of years. In typical beginner and intermediate Spanish classes, the periphrastic future tense (\textit{ir a} + infinitive) is introduced early and reinforced over the course of learning as the most common option to express future events. When object pronouns are introduced, the periphrastic future tense (with the matrix verb conjugated in the present tense) is one of the exemplary constructions (along with \textit{poder}, \textit{desear}, \textit{querer}, \textit{necesitar}) used to illustrate the optional word orders available for object pronouns with complex infinitival predicates. As such, late L2 learners of Spanish have a considerable amount of explicit and reinforced input (and usually an opportunity to practice) using the two word orders with the future tense of \textit{ir a}, conjugated in the present tense \textit{va}.

Second, a quantitative analysis of adult monolingual speech (the parents) from the Ornat-López file (CHILDES, MacWhinney, 2000) in Thomas (2007) pointed to a greater proportion of \textit{ir a} + \textit{infinitive} being used to express future events with the present tense form of \textit{ir}, such as \textit{va}. In these cases, monolinguals exhibited very high rates of proclisis in production (>90%). The form \textit{clitic} + \textit{va} (or other present tense forms of \textit{ir}) is likely an expression heard quite frequently by L2 speakers (teachers, television, radio, etc.). In this way, by an advanced level of proficiency, L2 speakers may have developed a strong representation for proclisis with \textit{va} + \textit{infinitive} based on their positive experience with this structure in input. On the other hand, using pronominal clitics with \textit{fue} + \textit{infinitive} is not likely to be as commonly heard, and as such, L2 speakers may have difficulty analyzing the marked word order (proclisis) when presented with marked and unmarked structural options. In hearing pre-verbal clitics with \textit{va}, L2 speakers appear able to easily analyze these forms; however, when they hear pre-verbal clitics with the verb form \textit{fue}, they appear to have more difficulty analyzing (and accepting) the structure. Despite having more difficulty with the marked structural option (proclisis) with \textit{fue} as compared to \textit{va}, L2 speakers are able to easily analyze the unmarked structure (enclisis). Future work to gauge the validity of a processing account of this differential treatment of \textit{va} and \textit{fue} with proclisis should include a reaction-time experiment. If L2 speakers are judging word order variants differently on the basis of the form of the matrix verb (as described above), then it should take them longer to analyze proclisis with \textit{fue} as compared to 1) enclisis with \textit{fue} and 2) proclisis with \textit{va}. 
6.4 Summary of Results

Overall, here I have shown that at advanced levels of proficiency both English heritage and L2 speakers of Spanish perform like Spanish monolinguals on the organization and limitation of proclisis (and thus, of verb incorporation) with optional clitic-climbing structures on the basis of structural, lexical and semantic variables. I will discuss the apparent group differences seen on the Acceptability-Preference Task—the relatively higher rates of acceptance of proclisis with embedded negation by bilingual groups and the relatively higher rates of acceptance of proclisis with *va* as compared to *fue* by L2 speakers—in the discussion chapter to follow. The biggest group differences for this study were found in the production task. In general, the heritage speakers outperformed the monolinguals and L2 speakers for the overall use of proclisis with optional clitic-climbing structures and for each individual verb type tested. Nevertheless, monolinguals and L2 speakers showed the most sensitivity to lexical cues in limiting their use of proclisis (and thus, verb incorporation) across verb types in production. As such, the monolinguals and L2 speakers patterned alike with respect to not only their production tendencies for this domain, but also with respect to a task-based asymmetry; monolinguals and L2 speakers exhibited the most task-based difference in their limitation of proclisis across verb types than heritage speakers. While these results point to L2 speakers being the same as monolinguals for this domain—and significantly different from heritage speakers—the statistical results here may underestimate a monolingual-bilingual difference. Monolinguals have been shown to vary their relative use of proclisis by mode of production (and thus, register type): higher rates in less formal registers and lower rates in more formal registers. Bilinguals may have more restricted knowledge of the competing registers that monolinguals use to organize and limit the grammatical optionality related to clitic-climbing constructions, and as such, the L2 speakers may pattern with the monolinguals here because of the formal nature of the experimental setting. I will discuss these results further in Chapter 7 and will discuss further research that should be conducted in order to tease apart the nature of the differences (and lack of differences) seen in this study.
7 General Discussion and Conclusions

7.1 Introduction

The main objective of this study was to empirically examine how different groups of speakers of the same language organize and limit grammatical optionality. More specifically, the main aim here was to determine if monolingual and highly proficient heritage and L2 speakers of Spanish limit the word order optionality of clitic-climbing constructions according to the same language-internal factors (syntactic, lexical, semantic). In general, I outlined this objective so as to better understand the extent to which the nature of language exposure (single- vs. dual-language contexts) and/or the age of language exposure (native vs. non-native) has an impact on the way that different types of speakers use the language system to limit grammatical optionality. Here, I reanalyze the notion of complete acquisition in contexts of bilingualism, and I discuss the ways in which linguistic, psycholinguistic and sociolinguistic factors may contribute to the unique and dynamic ways that different types of speakers (monolinguals-bilinguals, child-adult) use the language system through their lifespan.

7.2 Knowledge and Use of Clitic-Climbing by Spanish Speakers

Before turning to consider how the results seen in this study can be interpreted related to linguistic and situational factors associated with the nature and age of exposure, let us review the research questions and hypotheses outlined for this study and see how the results compare to these predictions for bilingual knowledge and use of Spanish clitic-climbing. In all cases, I will not repeat the specific hypotheses in detail here, I will just identify them in terms of whether I predicted no difference, age effect (L1, 2L1 ≠ L2), or a bilingual effect (L1 ≠ 2L1, L2).

Recall that the first group of questions dealt with the syntactic aspects of Spanish speakers’ knowledge of clitic-climbing:

1) Do early and late bilinguals of Spanish show evidence (in a receptive task) of instantiating a subset of infinitive-selecting verbs that permit both unmarked (enclisis) and Spanish-specific (proclisis) structural variants similar to monolinguals?
2) Do early and late bilingual speakers of Spanish know that embedded negation acts as a categorical constraint blocking the language-specific structural option (verb incorporation, proclisis) available for clitic-climbing constructions?

In Table 7.1 below, a summary of these two questions is addressed in terms of the hypotheses and the results found in this study.

Table 7.1

*Summary of results testing syntactic knowledge of verb incorporation*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>Hypothesis</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural</td>
<td>1) 6 verb types (X):</td>
<td>1) no</td>
<td>1) no</td>
</tr>
<tr>
<td></td>
<td>-proclisis: Juan la X ver.</td>
<td>difference</td>
<td>difference</td>
</tr>
<tr>
<td></td>
<td>-enclisis: Juan X verla.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2) embedded negation (EN)</td>
<td>2) no</td>
<td>2) no</td>
</tr>
<tr>
<td></td>
<td>-no EN: Juan (lo) quisiera saber(lo)</td>
<td>difference</td>
<td>difference</td>
</tr>
<tr>
<td></td>
<td>-with EN: Juan (*lo) quisiera no saber(lo)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While it had been shown in previous work on Spanish clitic-climbing that *evitar* as a matrix verb does not tolerate the language-specific structural variant (verb incorporation, proclisis) for this domain (Strozer 1976), L2 speakers of Spanish accepted proclisis with *evitar* at significantly higher rates than monolinguals, and at relatively similar rates as the other matrix verbs tested in this study (that have been shown previously to permit optionality). Nevertheless, given that the monolinguals exhibited indeterminate judgments on the acceptability of proclisis with *evitar*, no strong claim will be made here regarding group differences. In general, monolingual, heritage and L2 Spanish speakers here instantiated a similar subset of infinitive-selecting verbs that permit a language-specific structure-building option (verb incorporation) along with the unmarked one.

In terms of testing the knowledge that bilingual speakers have of the blocking effect of embedded negation, the expected outcome was confirmed by the results: no group differences. Even though bilingual groups accepted proclisis with embedded negation at significantly higher rates than monolinguals, all groups identified that infinitival predicates with embedded negation
are structurally different that those without embedded negation (i.e. that they do not tolerate verb incorporation, proclisis). So, even though infinitival predicates with embedded negation are largely underrepresented in input, both early and late bilinguals are able to acquire the syntactic constraints that limit verb incorporation in Spanish.

In addition to the syntactic knowledge that Spanish speakers have of verb incorporation, and thus clitic-climbing, I also tested these groups on their knowledge of non-categorical lexical and semantic constraints associated with clitic-climbing structures. Recall that this second group of questions dealt with testing monolingual, heritage and L2 speakers’ use of proclisis in a production task (PET), their sensitivity to a lexical restriction (type of matrix verb) of the optionality in production and receptive tasks (PET and APT), and their sensitivity to a semantic limitation (interpretation of the matrix verb) of the optionality in a receptive task (APT). I posed the following questions in relation to the way that Spanish speakers may use lexical cues to organize the optional grammar and to resolve the structural competition in production:

1) Do bilinguals (advanced Spanish) use proclisis as much as monolinguals in production?

2) Do early and late bilinguals differ in their use of lexical cues to limit the optionality of Spanish clitic-climbing structures in production (compared to monolinguals and compared to each other)?

3) Do Spanish speakers (monolinguals, bilinguals) use lexical cues (matrix verb type) to organize and limit the optionality of their grammar (in a receptive task)?

4) Do monolinguals and bilinguals exhibit a task-based asymmetry on the limitation of grammatical optionality by lexical-based cues provided by the matrix verb of the complex predicate (production vs. receptive tasks)?

In Table 7.2, a summary of these questions is addressed in terms of the hypotheses and results from this study.
Table 7.2.

Summary of results for use of lexical constraints on limiting optionality in production and receptive tasks

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>Hypothesis</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lexical</strong></td>
<td>1) Production task</td>
<td>1) i. Bilingual effect</td>
<td>1) i. her-mono/L2</td>
</tr>
<tr>
<td></td>
<td>i. overall use of proclisis</td>
<td>ii. Age effect</td>
<td>ii. no difference</td>
</tr>
<tr>
<td></td>
<td>ii. use lexical cues to limit proclisis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1) i. <em>ir a &gt; querer &gt; preferir</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2) Receptive task</td>
<td>2) Age effect</td>
<td>2) no difference</td>
</tr>
<tr>
<td></td>
<td>- use lexical cues to limit proclisis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1) i. <em>ir a &gt; poder &gt; querer &gt; necesitar &gt; preferir &gt; evitar</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3) Receptive vs. Production task</td>
<td>3) Age effect</td>
<td>3) her-mono/L2</td>
</tr>
<tr>
<td></td>
<td>- task-based asymmetry</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1) i. <em>ir a &gt; querer &gt; preferir</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While a number of previous studies have shown that early and late bilinguals of Spanish exhibit an enclisis bias in production for this domain (Meijer & Fox Tree, 2003; Pérez-Leroux et al., 2011a), the results here do not confirm a monolingual-bilingual difference, but rather, like Montrul (2010), these results point to heritage speakers having the strongest proclisis bias of the three groups. To account for these results, we should consider a language-external variable that has been shown to affect monolingual speakers’ choice of word order in production, or rather, the effect of register differences: higher rates of proclisis in informal production (spontaneous speech) as compared to a more formal mode of production (writing). If heritage speakers have predominant exposure to spontaneous native speech and L2 speakers have predominant exposure to written and variable speech varieties, then both groups may not have sufficient knowledge of how register differences affect the distributional patterns of word order optionality for this domain. Each group may choose between structural options in production on the basis of the experience they have attesting to the distributional probabilities of proclisis in more informal (heritage speakers) and formal (L2 speakers) varieties, and as such, there may be a monolingual-bilingual difference here: monolinguals may be using a more formal register, as well as language-internal factors, to limit their use of proclisis in production; however, bilinguals may be predominantly applying language-internal limitations (type of matrix verb) in production because they do not have as expanded knowledge of the way that register differences affect the relative
use of proclisis (verb incorporation) in production. Given the formal nature of this task, the monolinguals choose appropriately given the context, and as such, appear to perform more like L2 speakers than heritage speakers.

In terms of the sensitivity that these groups of speakers have to the lexical choice of matrix verb in restricting their use of proclisis, all groups appear to make use of this lexical cues (type of matrix verb) to limit the application of verb incorporation (proclisis) in production. However, again, while monolinguals and L2 speakers appear to be performing most alike, individual results suggest that L2 speakers (unlike monolinguals and heritage speakers) exhibit an enclisis bias in production for all the verbs that they are tested on here. Among all groups, L2 speakers appear to be most affected by either the strong experience they have with the unmarked structural option cross-linguistically (enclisis), which may lead to a processing effect in production (higher activation on unmarked structure), or by the quality or mode of input to which they have been exposed (direct experience with enclisis more overall). Whether a processing or input-related issue, L2 speakers, as non-native speakers of Spanish, have production outcomes for this domain that are different than native speakers, who rely on more Spanish-specific knowledge of this domain (lexical type of matrix verb and proclisis) to guide their choices in production.

In terms of the way that these three groups used lexical cues (type of matrix verb) to organize the optionality in their grammar (receptively), the results here run contrary to expectations: there were no group differences. All groups appeared to be sensitive to the type of matrix verb as a means to limit proclisis in a receptive task (lexical cues are not relevant only in production). Consequently, given the heritage speakers’ strong proclisis bias in production (compared to the other two groups), this group appears to organize (receptively) and apply (production) verb incorporation in similar ways despite the type of task. Monolinguals and L2 speakers, therefore, appear to pattern similarly in their tendency to exhibit task-type asymmetries, where they tend to produce a lot less proclisis by verb type than they accept/prefer in a receptive task. However, even though the monolinguals appear to have as pronounced a task-based asymmetry as L2 speakers, this similarity between monolinguals and L2 speakers may be overestimated. As noted above, the rates of proclisis that monolinguals produce here are lower than previously reported for spontaneous speech (Davies, 1995; Thomas, 2007), an outcome most likely related to the
formal nature of the task. Consequently, these results may underestimate a monolingual-bilingual difference; more testing is required to assess this question.

Finally, to test how sensitive speakers are to semantic information for this domain, I pursued the following questions:

1) Do Spanish monolinguals use semantic cues to organize and limit grammatical optionality associated with Spanish clitic-climbing (receptively)?

2) Do Spanish-English bilinguals use semantic cues to organize and limit the grammatical optionality associated with Spanish clitic-climbing (receptively)?

3) How do Spanish-English bilinguals compare against Spanish monolinguals in their use of interpretational cues to limit the grammatical optionality associated with Spanish clitic-climbing (receptively)?

The semantic variable is summarized below in Table 7.3 with the hypotheses and the results for this variable.

Table 7.3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>Hypothesis</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic</td>
<td>1) Interpretation of ir a -future&lt;br&gt;Juan (la) <em>va a comprar</em>(la)&lt;br&gt;-motion&lt;br&gt;Juan (la) <em>fue a comprar</em>(la)</td>
<td>1) Age effect</td>
<td>1) Reverse age effect</td>
</tr>
</tbody>
</table>

I found here that neither the monolinguals nor the heritage speakers used semantic information related to the tense or situational aspects of the predicate to organize the optionality for this domain. These results run counter to the original hypothesis that predicted that native speakers would be sensitive to interpretational cues in limiting the acceptability of proclisis in a receptive task, unlike non-native speakers (L2 speakers). Here, L2 speakers did exhibit sensitivity for this particular variable: they accepted and preferred sentences with proclisis at higher rates when the verb *ir a* was conjugated in the present tense and presented in a future context than when it was conjugated in the perfective past tense and presented in a context of motion. While
these results may suggest that non-native speakers are using an interpretational strategy to limit the optionality that native speakers don’t use, the more plausible explanation of these results is that the L2 speakers are more comfortable accepting proclisis with the present tense form _va_ as compared to the past tense _fue_. L2 speakers may have had strong positive evidence for ‘_lo va a X_’ in their experience with Spanish (explicitly and implicitly) as compared to ‘_lo fue a X_’. As such, they may be most confident in their judgments of the first form as compared to the second.

In sum, in terms of the syntactic knowledge that monolinguals, heritage and L2 speakers of Spanish have of optional clitic-climbing (i.e. infinitival predication), it appears as though neither one’s experience with dual language exposure, nor a late age of language exposure limits a speaker’s ability to

i) acquire verb incorporation as a structural option for infinitival predication (with pronominal objects)
ii) establish a basic subset of infinitive-selecting verbs that can undergo verb incorporation with an infinitive when pronominal objects are selected
iii) know when verb incorporation is syntactically restricted.

All groups appear to use the type of matrix verb (lexical factor) as a way to constrain the optionality, both in the organization of the grammar (receptively) and for choosing between two structures/forms in production. Nevertheless, a closer examination of the results for the lexical variable points to bilingual speakers performing in a significantly different way in production. This may be so because they have different experiences with exposure to Spanish: heritage speakers are mostly exposed to spontaneous native speech and interactions in a Spanish vernacular, and L2 speakers have strong experience with writing, explicit grammatical instruction and a number of speech varieties (more corrected/formal speech, native and non-native speech). Heritage speakers use proclisis at the highest rates of all speakers and L2 speakers use proclisis at the lowest rates here, but within the range of variability permitted by monolingual speakers. However, individual results suggest that, while L2 speakers almost always exhibit an enclisis bias regardless of verb type, monolinguals exhibit more variability in their word order choices overall, and are guided more by verb type in their word order biases for this domain. As such, the production tendencies of each group are rather unique here. Although more testing is required to confirm the nature of these tendencies (whether psycholinguistically or sociolinguistically
motivated), these results suggest that what is complete for any group of speakers here has to do with a unique experience based on the age of exposure to Spanish.

Finally, the native Spanish speakers tested here do not appear to use semantic cues provided by the tense of the matrix verb nor the predicate in judging sentences with proclisis. L2 speakers appear more comfortable in allowing proclisis with the present tense form of ir a as compared to the perfective past tense, a result that may be reflective of L2 speakers’ strong positive experience with the structure clitic + va as compared to clitic + fue. Whether any group uses semantic constraints (with the same matrix verb) to resolve competing structural options in production is a question not addressed here.

7.3 Age of Exposure: Language and Cognition

This study confirms that English speakers are able to acquire both the clitic status of object pronouns and the language-specific feature (verb incorporation) available for complex predicate formation in Spanish regardless of their age of exposure to the language. These results are similar to previous studies that have examined the acquisition of pronominal clitics with simple and complex predicates in Spanish by early and late Spanish-English bilinguals (Duffield & White, 1999; Montrul, 2010). Adult L2 speakers do not appear to be limited by some critical or sensitive period in accessing linguistic information not characteristic of their L1 to constrain their late-acquired grammar. Within the generative framework, these results generally support a full access approach to the availability of a universal inventory of language information (UG) to build the linguistic knowledge of a late-acquired L2 (i.e. Schwartz & Sprouse, 1996), and go against theories that propose a critical period for L2 acquisition (i.e. Bley-Vroman, 1989).

Where L2 speakers appear to have more difficulty, however, is in providing reliable judgments for structures that are underrepresented in Spanish input. While L2 speakers know the difference between predicates with and without embedded negation (and thus, show evidence of knowing the syntactic constraints of verb incorporation), and that proclisis is freely allowed regardless of the tense of the verb ir a (to be going to), they tend to perform in a more variable way for these domains than monolinguals, which makes it appear as though their knowledge of this domain may be just short of complete. Nevertheless, others examining L2 speakers (and bilingual speakers) in empirical studies have pointed out that this relative experimental
uncertainty is common for non-native speakers, especially when they have to make judgments on the ungrammaticality of certain sentences (Schutze 1996). Further, early bilinguals also exhibited this non-monolingual-like behaviour for sentences with embedded negation. The failure of non-native speakers to behave in a monolingual-like way on the judgment of these ungrammatical sentences, therefore, appears to be more associated with the speakers’ status as bilinguals, rather than as non-native (late adult) speakers of Spanish. As such, the lack of complete behaviour here exhibited by L2 speakers (if complete is associated with monolinguals) could not be due to some cognitive deficit associated with acquiring a language later in life (as an adult) since early bilinguals perform the same way.

Further, although the production results reported here only quantify the pronominal responses by the participants in this study, I will make a brief comment about the other response types, null objects and overt lexical objects. All groups of speakers produced less than 10% of null objects overall for all verb types. When null objects were used, speakers used them consistently in target-like ways with prototypical verb pairings, as in (81):

(81)  Q. ¿Qué va a hacer cada uno de ellos con los platos? “What is each of them going to do with the plates?”
A. Marta va a lavar Ø y Pedro va a secar Ø. “Martha is going to wash Ø and Peter is going to dry Ø.”

These results are generally consistent with the fact that the participants tested in this study were advanced speakers of Spanish, and as such, should have known that the responses of this particular task typically required the overt expression of a direct object.

Where the monolingual, heritage and L2 speakers of Spanish do appear to diverge somewhat is in the use of overt lexical objects. Among the four response types evaluated for the production task (proclisis, enclisis, overt lexical objects and null objects), the use of enclisis and overt lexical objects represent responses where speakers used the unmarked structure-building option available for this domain. Of all groups, the L2 speakers used the unmarked structure (enclisis + DP) at the highest rates for the frequently-used verbs *ir a* (69.7%) and *querer* (82.5%). Heritage speakers and monolinguals used the unmarked structure relatively less for *ir a* (26.6% and 61.4% respectively) and *querer* (24.1% and 71.8% respectively). Below I consider these
results by group to determine the strategies that the different speakers employ in making structural choices at the moment of production.

Overall, monolinguals were relatively consistent in their response types: low rates of object omissions, around 20% lexical objects (DPs) and the remaining proportion divided variably by verb type between proclisis and enclisis. Note that monolinguals appear to be strongly guided by lexical type in their use of verb incorporation (proclisis), but use the unmarked structure-building option (enclisis + DP) proportionally more in this task for all verbs than the Spanish-specific option, verb incorporation (proclisis). As mentioned previously, these results are inconsistent with previous studies that have reported higher rates of proclisis for these verb types in the spontaneous speech of Spanish monolinguals (Davies 1995; Thomas 2007). Although I did not collect enough spontaneous speech data from the monolingual participants in this study to evaluate clitic-climbing structures, I would suspect that these relatively lower rates of proclisis by monolinguals is due to the formal nature of the task, and perhaps due to the use of a more corrected register by the (educated) monolingual participants who would have identified their tester (me) as a non-native Spanish speaker.

Conversely, heritage speakers and L2 speakers don’t appear to be guided as strongly by the type of matrix verb in their use of verb incorporation. Heritage speakers treated ir a and querer alike and used verb incorporation at proportionally greater rates (>70%) than the unmarked structure (enclisis + DP) for these verbs. For preferir, the heritage speakers went in the expected direction of using less verb incorporation, but still used it at higher rates (23%) than the monolinguals (0%) and L2 speakers (4%). Similar to the heritage speakers, but at the other end of the spectrum, L2 speakers used verb incorporation at relatively low rates (<30%) for the verbs ir a and querer, but then performed most like monolinguals for the verb preferir (unlike heritage speakers). Here, for all verbs, L2 speakers used the unmarked structural option (enclisis + DP) at rates of 70% or more while heritage speakers, at least for the verbs ir a and querer, used the unmarked structure (enclisis + DP) at rates under 30%. Clearly the way that these two groups of bilinguals resolve the optionality in production is different. In this section, I consider an age-related processing explanation of these results, where L2 speakers exhibit an enclisis bias overall as compared to heritage speakers of Spanish.
Recall the language system outlined for bilingual speakers in Chapter 3: one lexicon with lexical entries organized as a network and an invariant computational system that can analyze or build structures as long as they conform to rules on structure-building typical of natural language. How any lexical items (substantive or non-substantive categories) would get selected and passed onto the computational system in the production of a single-language utterance has to do with the relative activation levels of those items, determined by both procedural and declarative processes associated with the semantics/pragmatics/discourse context of the expression to be articulated and the relative fluency of the speaker (overall and in comparison to the non-target-language). As such, the activation of the lexicon is language-non-specific and language selectivity is determined by two main factors: 1) the ability of the speaker to use executive control function to inhibit the activation of entries from the non-target-language (Bialystok, 1987; 2009); and 2) the relative proficiency, or lexical robustness of the speaker as defined by their size of lexicon and speed of lexical access for the target language; those with greater proficiency/lexical robustness, and thus, a larger inventory of target-language features and faster lexical access, would rely less on the function of executive control to inhibit the non-target-language resources, and could instead select from the intended language directly without the interference of the non-target-language nor the cognitive load implied in the use of executive control (Costa & Santesteban, 2004; Schweiter & Sunderman, 2008).

Nevertheless, the idea of an autonomous selection process in bilingualism is unlikely as strict language selectivity is often incompatible with the notion of an efficient language system. For example, if two languages use the same (perhaps unmarked) structure-building process to construct a linguistic expression with the same conceptual plan, then it would be inefficient for such a system to instantiate language-specific rules for this domain; hence, the notion of unmarked structures. Bilingual priming studies, such as the one by Meijer & Fox Tree (2003) for Spanish-English bilinguals, point precisely in this direction: that bilinguals pool syntactic resources that are used commonly for both languages. For Spanish and English infinitival predicates with pronominal objects, the unmarked (or common) structure is to build an infinitival predicate grammatically in a bottom-up fashion: argument to infinitive, infinitival predicate to matrix verb. As such, if the activation of the bilingual lexicon is non-language-specific, then the invariant and unmarked structure building process used for infinitival predication cross-linguistically may have a stronger activation in those with more robust English experience, such
as L2 speakers. Even though these speakers can acquire and use verb incorporation for infinitival predication, they tend to do so less than native speakers (heritage speakers and monolinguals) regardless of their relative fluency in Spanish (recall 0% proclisis in narrative task for low-intermediate L2 speakers in Montrul, 2010).

Conversely, heritage speakers, from low to advanced proficiency levels, tend to use proclisis at the highest rates in production, including at higher rates than monolingual speakers (recall low-intermediate heritage speakers outperformed native controls on use of proclisis in narrative task in Montrul, 2010). It appears as though heritage speakers, even those who are English-dominant, are able to use a (more marked) Spanish structure-building process that requires good procedural control over target-language lexical and grammatical processes. The heritage speakers’ early exposure to Spanish, and relatively greater communicative use/interaction as children (as compared to adolescence and adulthood), may have helped these speakers to develop the automatization necessary (direct use of language-specific resources) to apply the process of verb incorporation without searching the system for a way to resolve grammatical optionality for this domain. In addition, this process does not appear to erode as the speaker gets older and becomes more dominant in a language (English) that offers strong evidence of the unmarked structure-building option for this domain. As such, this area of grammar appears to support the Regression Hypothesis of Jakobson (1941, cf. Montrul, 2008); the more early-acquired processes/features are less prone to attrition (first in, last out).

Again, here is where the L2 speakers may be at a disadvantage as compared to heritage speakers as a result of their age of exposure to Spanish. The process of verb incorporation, as a lexical process, would require relatively fast (and automatic) target-language-specific processing capabilities and relatively little use of declarative resources in the production process. Even though adult heritage speakers of Spanish may be more dominant in English (dominance defined in terms of linguistic and sociolinguistic abilities), they would have acquired the process of verb incorporation as an early-acquired feature, and would have continued to attest to its robust use through mostly informal (vernacular) and conversational interactions with other native Spanish speakers. Conversely, L2 speakers would have acquired verb incorporation after instantiating a grammar (their L1) that invariably used the unmarked structure-building option. As such, these speakers may not have the required speed of processing to use verb incorporation (because they
didn’t necessarily need it—the unmarked structure is a well-formed way to express pronouns with infinitival predicates in Spanish), or may be using too much declarative knowledge (thinking too much about the expression they want to formulate) to be able to overcome the relatively high activation levels of the unmarked option as easily as heritage speakers. As such, unlike heritage speakers, who may have to use little attentional control to use verb incorporation (it is a more automatized process for this group of speakers), L2 speakers may have to use more cognitive resources (i.e. executive control) to overcome the strong activation of unmarked structural variant. Here, English does not appear to have a direct influence on the choices that L2 speakers make in production, but may indirectly increase the relative activation of the feature associated with unmarked structure-building, and as such, result in a quantitative difference between the use of proclisis by native and non-native Spanish speakers. Nevertheless, without testing other language-pairings, we can not know for certain if the tendency of L2 speakers to use the more unmarked structure-building option is the result of indirect influence from English (cross-language influence) or is the result of the speaker defaulting to the unmarked structure because of his or her lack of ability to resolve the optionality in the most language-specific way (general bilingual effect).

7.4 Age of Exposure: Context of Exposure

In the previous section, I explained the relative enclisis bias exhibited in production by advanced L2 speakers of Spanish as the result of this group’s relatively greater reliance on the unmarked structure-building process to resolve the optionality of this grammatical domain, as compared to Spanish native speakers (heritage speakers and monolinguals). In this discussion, I alluded to the fact that heritage speakers and L2 speakers would have had different contexts of exposure related to their age of exposure to Spanish: for heritage speakers, early exposure to naturalistic and informal speech via input/interactions with other native Spanish speakers, and for L2 speakers, early predominance of written and formal/grammatical exposure with increasing interactions in the target language with both native and non-native speakers. The results above, where heritage speakers and L2 speakers resolve the optionality associated with Spanish clitic-climbing in different ways, may also be explained, therefore, according to experiential factors associated with these different contexts of exposure to Spanish.
First, if L2 speakers only have more formal/academic experiences with Spanish and interact with mainly other bilinguals, then they may never be exposed to naturalistic input with the (higher) rates of proclisis, such as identified in Davies (1995) for monolingual Spanish speakers in spontaneous speech. Instead, the relatively lower rates of proclisis (for all verb types) used by L2 speakers may be the result of their primary experience with proclisis (for all verb types) in more formal contexts and in interactions with other bilingual speakers. As others have shown empirically (Paradis & Navarro, 2003; Hulk & Cornips, 2006; Rothman, 2007), a different quality of input, and lower probabilities associated with the use certain options, may serve as the input for L2 (or bilingual) speakers, who will produce a target language (rates of proclisis overall and by verb type) that reflects this experience. Heritage speakers, on the other hand, would have more restricted experiences with input from mainly spontaneous and informal speech varieties; consequently, their primary source of input could exhibit relatively high rates of proclisis (higher than reported in any study), and as such, they too would produce a target language that reflects this experience. Nevertheless, as other studies have begun to show, the language used by many Spanish heritage speakers in North America (and by other bilinguals in language contact situations) is a variety unique to these types of bilingual speakers; as such, as these speakers begin to have children and speak to them in Spanish, what appears to be an incomplete grammar may, in fact, be a unique grammar that bilinguals have developed on the basis of their overall language experience in language contact situations (i.e. Sánchez, 2004; Bullock & Toribio, 2004). In these cases, complete acquisition for a bilingual (from a linguistic point of view) is different than what could ever be expected in an idealized monolingual environment.

The long-term effects of these more restricted experiences (sociolinguistically) or different input types (linguistically) for the domain of clitic-climbing, may be that these bilingual groups are not able to adjust their relative use of proclisis to what is appropriate to different contexts of use, such as in different modes of production (speech vs. writing) or according to different language registers (formal vs. informal production). The results from this study suggest that this is so for heritage speakers: despite performing in a relatively formal task, these speakers use significantly higher rates of proclisis than monolingual speakers, who appear to adjust their speech register, and as such, their relative use of verb incorporation. Nevertheless, given that no sociolinguistic variable was explicitly tested in this study, the sociolinguistic question remains open at this time (for the monolinguals too). An interesting follow-up here would be to test
monolinguals, heritage and L2 speakers of Spanish with this protocol, record them in more naturalistic, spontaneous conversations and test them on their relative word order preferences for clitic-climbing structures in a written mode of production. In this way, we could see if all Spanish speakers use the same language-external resources to limit the optionality for this domain. The notion of what complete acquisition means for this domain, thus, remains open as well.

7.5 Optionality and Variability in Bilingualism: Concluding Remarks

In this study, I examined how different groups of Spanish speakers organize and limit core grammatical optionality—Spanish clitic-climbing—according to their age and context of exposure to the language. Through the review of previous empirical work and through the results presented in this study, I found that like monolinguals, bilinguals tend to limit optional grammars, but do so in different ways according to the unique cognitive and experiential factors associated with their age of exposure to bilingualism. Spanish-English bilinguals raised in North America tend to have little formal education in Spanish, and as such, tend to limit their grammar in the same way that other native speakers would in casual conversation: for clitic-climbing they use high rates of proclisis (the marked Spanish-specific structure) even when monolinguals do not. Adult L2 speakers, on the other hand, have relatively less informal interaction in Spanish (compared to monolinguals and heritage speakers), and thus, appear to rely more on unmarked linguistic strategies permitted for clitic-climbing structures in Spanish, much in the same way that monolinguals perform under these experimental conditions. In production, therefore, there appears to be both an age effect (L1, 2L1 ≠ L2) and a bilingual effect (L1 ≠ 2L1, L2).

In the first place early bilinguals (native speakers) perform differently than late bilinguals (non-native speakers); this difference may be due to both processing and experiential-related factors associated with age of exposure to Spanish. Early (heritage) bilinguals would have acquired verb incorporation, and thus proclisis, on the basis of native, naturalistic input before they began primary school and their dominant exposure patterns switched to English. The processing abilities required to use verb incorporation in speech don’t appear to erode over time and in the face of shifting language exposure; consequently, even when English is the dominant language of these heritage speakers, they still appear to retain a good ability to apply verb incorporation when speaking in Spanish. Late bilinguals (adult L2 speakers of Spanish) on the
other hand, must overcome strong experiences with an unmarked structure in order to apply verb incorporation in production. Given that the unmarked structure is totally well-formed for cases of optional clitic-climbing, the strategy used by most L2 speakers, even at advanced levels of proficiency, is to use the unmarked structure (either with a pronoun or a lexical object). Again, this tendency toward an enclisis bias may be the result of both cognitive- and experiential-related factors. Further, it may be more of a processing burden for L2 speakers to apply the process of verb incorporation in oral production (on-line) because this implies relatively automatic language-specific processing, as well a good ability to inhibit the non-target-language (their L1 in this case). Further, because their context of exposure is strongly represented by written, academic and bilingual sources of input, late bilinguals may be performing in a way that directly reflects this experience: with less proclisis overall.

Secondly, monolinguals perform differently than bilinguals in their apparent ability to vary their rates of proclisis according to factors associated with mode of production, and perhaps the formality of the discursive context. The results collected here suggest that this is not necessarily so for heritage and L2 speakers of Spanish (in North America), especially for the heritage speakers who use much higher rates of proclisis than the monolinguals tested here. As such, the more restricted experiences that heritage and L2 speakers have, because of the nature of their language exposure (as bilinguals in North America), may restrict them from acquiring the multiple registers that most monolingual speakers have. Consequently, these bilinguals may always be English-dominant” (both linguistically and sociolinguistically) because of their relatively unequal sociolinguistic abilities in English and Spanish (the result of being primarily educated in English). Given the results of the receptive task, heritage and L2 speakers have no trouble in acquiring the linguistic constraints on verb incorporation in the same way as monolinguals, even if they are less determinate in their judgments for ungrammatical and not commonly attested structures. As such, the bilinguals exhibit so-called complete acquisition for this domain from a linguistic point of view. However, they may not always sound/perform as monolingual-like for every discursive context, and as such, one might assume they lack complete acquisition. Here, I would like to suggest that they have indeed completely acquired this domain given their context of exposure and unique communicative needs (i.e. communicating with other bilinguals, etc.).
Finally, given that clitic-climbing represents core optionality from a grammatical point of view, one would expect a considerable amount of variability to emerge in any group’s performance, either receptively or productively. This is indeed what we see in this study, especially for areas where speakers are somewhat unsure how to judge not-so-grammatical structures. As such, monolinguals exhibit indeterminate judgments for proclisis with the verb *evitar* and bilinguals exhibit indeterminate judgments for proclisis with sentences with embedded negation. Persistent variability in advanced grammars of bilinguals, therefore, tends to arise where there is some learnability issue; where a particular form may be underrepresented/ambiguous in input, and thus, the speaker has to rely on overall language knowledge to resolve the ambiguity/difficulty. Monolinguals, therefore, are not immune to such variability, just that they tend to not exhibit it as often because they can only rely on the linguistic resources of the target language to resolve the learnability issue (and tend to do so by adulthood). Bilinguals have more expanded language resources at their fingertips, and as such, if cross-language and/or universal language resources can provide a satisfactory analysis for a particular problem, then that speaker may use it in development and/or persist in its use as an adult because they don’t recognize/don’t care to recognize that their analysis/output is non-target-like. Over time, therefore, further generations of bilingual speakers in North America may come to acquire a Spanish grammar that is already bilingual, and as such, the notion of complete acquisition for these speakers must be determined on a generation by generation, or group by group basis.

### 7.6 Direction for Future Research

This investigation could be expanded to answer further linguistic questions, and begin to examine the psycholinguistic and sociolinguistic questions left open in this study. In the first place, in addition to testing more participants per group, I would like to further investigate the role that semantics plays in the way that Spanish speakers limit the optionality for this grammatical domain; more specifically in the way that native speakers may use verb incorporation (productively) depending on the relative interpretation of the same matrix verb, such as with frequent modal and desiderative expressions such as *poder* (to be able to) and *querer* (to want to). The goal here is to see whether semantics is a factor use more in production as a way to resolve the optionality for this domain even if they don’t use it to organize the optionality receptively (here they appear to use lexical cues more). Further, I would like to conduct a correlational study
mapping the individual results here to both the non-linguistic and linguistic measures related to the participants’ linguistic experience and relative level of proficiency. Which factors are most likely to be correlated to a speaker’s more reliable use of proclisis (verb incorporation, a more automatized language system) in production?

Additionally, I would like to develop some on-line tests to determine if L2 speakers employ more cognitive resources in the analysis and/or use of clitic-climbing constructions in Spanish as compared to heritage speakers. Here, I plan to test speakers of different ages and proficiency levels to determine if the relative shifting balance of proficiency in Spanish has an effect on how efficiently different types of bilinguals (heritage/L2 speakers) are able process verb incorporation receptively and productively. The psycholinguistic work may help us to better define the role that age as a cognitive variable plays in persistent variability in bilingual grammars. Finally, it would also be useful to test Spanish bilinguals of all proficiency levels (as adults) on the way they handle the structural optionality of Spanish clitic-climbing according to a variety of language-external factors, including mode of production (speech and writing) and different speech registers. This work may help us to better identify the communicative gaps that bilinguals may have (writing skills, conversational skills), that in turn may help us to better define the educational needs of these Spanish speakers in secondary and post-secondary levels of study.
References


