Lexical Availability in Diaspora Spanish:
A Cross-generational Analysis of Chilean Swedes

BY
MARYANN NEILSON PARADA
B.A., Brigham Young University, 2006
M.A., Brigham Young University, 2008

THESIS
Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Hispanic Studies
in the Graduate College of the
University of Illinois at Chicago, 2016

Chicago, Illinois

Defense Committee:
Kimberly Potowski, Chair and Advisor
Richard Cameron, Linguistics
Kara Morgan-Short, Hispanic Linguistics
Giedrius Subacius, Lithuanian Studies
Marta Fairclough, University of Houston
For my children, Marcus, Antonio, and Evelina. May you find joy in pursuing your dreams, and in making life better for others.
ACKNOWLEDGEMENTS

I would like to thank my thesis committee—Kim Potowski (Chair), Richard Cameron, Kara Morgan-Short, Giedrius Subacius, and Marta Fairclough—for their mentorship and insightful feedback. I am fortunate to have benefited from their scholarly expertise and friendship. I am also grateful for the support I received through the UIC Chancellor's Graduate Research Fellowship and the Provost's Award.

I am also indebted to all of the individuals of the Chilean community in Stockholm who served as study participants. It was a truly rewarding experience to interact with them and learn from their experiences. I thank them for their generosity.

I extend my thanks to Orlando Alba, Robert Smead, and Scott Alvord of Brigham Young University under whose guidance the first ideas for this project emerged. My thanks also go to Jeffrey Turley of Brigham Young University for advising me to "go as far as I could". I also thank Yasmina Vallejos for her helpful responses to my many inquiries regarding aspects of Chilean Spanish. I also owe my gratitude to my heritage learner students in Chicago, who have been a constant source of inspiration.

I thank my fellow students and colleagues, especially Lillian Gorman, whose bright minds kept the bar very high. I also thank my faith community and friends for their voices and actions of support throughout this endeavor. Specifically, I thank Vivian Malli, Ravi Malli, Anya Bennee, Marjorie Meservey, and Charlotte Tamanaha. I am also deeply grateful to Eva Maxwell and the Gándara family for loving and caring for my children.

Lastly, but not least, I would like to thank my husband, Andrés, for his unending encouragement, inspiration and countless other forms of support. My success would have been
impossible without him. I thank my siblings for their loyal support and words of wisdom, as well as my dear parents, who always encouraged me to pursue education.
**TABLE OF CONTENTS**

1. INTRODUCTION .................................................................................................................. 1
   1.1 Introduction ....................................................................................................................... 1
   1.2 Statement of the Problem ............................................................................................... 4
       1.2.1 Heritage language acquisition and the lexicon ......................................................... 4
       1.2.2 Qualitative lexical change ...................................................................................... 10
       1.2.3 Goals of the study ................................................................................................. 16
2. LITERATURE REVIEW ......................................................................................................... 20
   2.1 Heritage Speakers ........................................................................................................... 20
   2.2 The Chilean Diaspora and the Chilean Minority in Sweden ............................................. 23
       2.2.1 Transmigration and the national, cultural, and linguistic identification of Chilean- 
            Swedes ....................................................................................................................... 28
       2.2.2 Maintenance, attrition, and Spanish-Swedish contact .............................................. 29
   2.3 Trilingual Heritage Speakers ......................................................................................... 33
   2.4 Spanish Heritage Speaker Lexicon ................................................................................ 38
       2.4.1 Defining vocabulary and lexical knowledge ............................................................. 38
       2.4.2 The SHS lexical repertoire ...................................................................................... 43
       2.4.3 Research in SHS lexicon ......................................................................................... 44
   2.5 Lexical Availability ......................................................................................................... 57
       2.5.1 Lexical availability research in Spanish ................................................................. 61
   2.6 Social and Language Experience Factors ..................................................................... 66
       2.6.1 Exile background ...................................................................................................... 67
       2.6.2 Birth order ............................................................................................................... 69
       2.6.3 Gender ..................................................................................................................... 71
       2.6.4 Parental education .................................................................................................. 73
       2.6.5 Heritage language instruction ................................................................................. 74
       2.6.6 Proficiency measures and language use/exposure .................................................. 76
   2.7 Language Change and the Lexicon ............................................................................... 76
       2.7.1 Language contact and the lexicon ........................................................................... 77
       2.7.2 Dialect contact ......................................................................................................... 83
       2.7.3 Cultural linguistics and the lexicon ........................................................................ 88
   2.8 Conclusion .................................................................................................................... 92
3. METHODOLOGY ................................................................................................................... 94
   3.1 Introduction .................................................................................................................... 94
   3.2 Participants ..................................................................................................................... 95
   3.3 Data Collection Procedure ........................................................................................... 96
       3.3.1 Questionnaire ........................................................................................................... 96
       3.3.2 The lexical availability task .................................................................................... 100
3.4 Processing and Coding.................................................................................................................. 105

4. RESULTS AND DISCUSSION............................................................................................................... 112
  4.1 Research Question 1a: Quantitative Comparisons................................................................. 112
  4.2 Research Questions 1b and 1c: Social and Language Factors ........................................... 124
  4.3 Research Question 2a: Lexical Compatibility ............................................................................ 131
  4.4 Research Questions 2b: Contact Lexicon ................................................................................. 145
     4.4.1 English contact lexicon ...................................................................................................... 145
     4.4.2 Swedish contact lexicon .................................................................................................... 149
  4.5 Research Question 2c: Dialect Contact (Peninsularisms) ....................................................... 153
  4.6 Research Question 2d: Conceptual Patterns ............................................................................ 155

5. CONCLUSIONS.................................................................................................................................... 168
  5.1 Primary Research Conclusions .................................................................................................. 168
  5.2 Study Significance ....................................................................................................................... 174
  5.3 Study Limitations ......................................................................................................................... 176
  5.4 Future Directions ........................................................................................................................ 178

REFERENCES ........................................................................................................................................ 181

APPENDICES ......................................................................................................................................... 202
  APPENDIX A: Questionnaire .......................................................................................................... 202
  APPENDIX B: Diplomas of Spanish as a Foreign Language (DELE) ............................................ 207
  APPENDIX C: Controlled Vocabulary Levels Productive Abilities Test .................................... 213
  APPENDIX D: Association Task ....................................................................................................... 215

VITA .................................................................................................................................................. 216
LIST OF TABLES

2-1: TOP TEN COUNTRIES WITH THE LARGEST CHILEAN POPULATIONS, BOTH CHILEAN-BORN AND FOREIGN-BORN ................................................................. 24
2-2: FOREIGN-BORN SPANISH SPEAKING POPULATION IN SWEDEN BY COUNTRY OF ORIGIN 1900-2014 ............................................................... 26
2-3: SEMANTIC PROMPTS USED IN MORENO-FERNÁNDEZ (2007) .................. 49
2-4: BORROWING TYPOLOGY ........................................................................... 80
3-1: SUMMARY OF G2 SOCIAL VARIABLES .................................................. 97
3-2: SUMMARY OF LANGUAGE VARIABLES ..................................................... 98
3-3: ENGLISH CONTROLLED PRODUCTIVE ABILITY TEST .......................... 100
3-4: SEMANTIC DOMAINS AND CATEGORIES ................................................. 103
3-5: LEXICAL AVAILABILITY EDITING GUIDELINES ........................................ 105
3-6: SUMMARY OF RESEARCH QUESTIONS AND CORRESPONDING ANALYSES... 111
4-1: DOMAIN RANKING BY SPEAKER GROUP ACCORDING TO MEAN NUMBER OF WORDS ........................................................................................................ 112
4-2: PAIRWISE COMPARISONS OF GROUP MEANS ........................................ 116
4-3: BETWEEN-GROUP SUPRA-CATEGORY MEAN COMPARISONS .................. 119
4-4: COMPARISON OF MEANS WITH HOMELAND SPEAKERS .......................... 121
4-5: COMPARISON OF MEANS WITH U.S. HERITAGE SPEAKERS ................... 123
4-6: G2 PAIRWISE COMPARISONS FOR SOCIAL VARIABLES ......................... 124
4-7: SUMMARY MEANS OF LANGUAGE VARIABLES ........................................ 126
4-8: SELF-REPORTED PROFICIENCY ACROSS LANGUAGES ............................. 128
4-9: CORRELATIONS BETWEEN LANGUAGE VARIABLES AND TOTAL WORDS .... 130
4-10: OVERVIEW OF INTERGENERATIONAL LEXICAL COMPATIBILITY IN SIX DOMAINS ......................................................................................................... 132
4-11: GIRARUDDS INDEXES IN COMPARISON ................................................. 133
4-12: WORDS IN COMPARISON (HUMAN BODY) ........................................... 135
4-13: WORDS IN COMPARISON (FURNITURE) ............................................... 138
4-14: WORDS IN COMPARISON (MODES OF TRANSPORTATION) ................. 140
4-15: WORDS IN COMPARISON (POLITICS) .................................................... 143
4-16: RATE AND DISTRIBUTION OF SWEDISH CONTACT LEXICON ............... 150
4-17: PENINSULARISMS .................................................................................. 154
4-18: TOP 15 WORDS IN COMPARISON (HUMAN BODY) .............................. 156
4-19: TOP 15 WORDS IN COMPARISON (FURNITURE) ................................. 157
4-20: TOP 15 WORDS IN COMPARISON (MODES OF TRANSPORTATION) ...... 159
4-21: TOP 15 WORDS IN COMPARISON (PROFESSIONS AND TRADES) ......... 161
4-22: TOP 15 WORDS IN COMPARISON (ECONOMY) ...................................... 163
4-23: TOP 15 WORDS IN COMPARISON (POLITICS) ........................................ 165
LIST OF FIGURES

Figure 3-1. Sample lexical availability output list................................................................. 107
Figure 4-1. Group means along the domain continuum.......................................................... 114
Figure 4-2. G1 vs. G2 mean differences .................................................................................. 115
Figure 4-3. Sample G2-G1 divergent/convergent lexicon (Human Body) ......................... 136
Figure 4-4. Sample G2-G1 divergent/convergent lexicon (Furniture)................................. 139
Figure 4-5. Sample G2-G1 divergent/convergent lexicon (Modes of Transportation) .... 142
Figure 4-6. Sample G2-G1 divergent/convergent lexicon (Politics) ...................................... 144
Figure 4-7. Relative prominence of top 15 ranked words (Human Body) ......................... 157
Figure 4-8. Relative prominence of top 15 ranked words (Furniture) ............................... 159
Figure 4-9. Relative prominence of top 15 ranked words (Modes of Transportation) ...... 160
Figure 4-10. Relative prominence of top 15 ranked words (Professions and Trades) ....... 162
Figure 4-11. Relative prominence of top 15 ranked words (Economy) ............................. 164
Figure 4-12. Relative prominence of top 15 ranked words (Politics) ................................. 166
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Availability Index</td>
</tr>
<tr>
<td>G1</td>
<td>First Generation</td>
</tr>
<tr>
<td>G2</td>
<td>Second Generation</td>
</tr>
<tr>
<td>HL</td>
<td>Heritage Language</td>
</tr>
<tr>
<td>HS</td>
<td>Heritage Speaker</td>
</tr>
<tr>
<td>L1</td>
<td>First Language</td>
</tr>
<tr>
<td>L2</td>
<td>Second Language</td>
</tr>
<tr>
<td>L3</td>
<td>Third Language</td>
</tr>
<tr>
<td>RQ</td>
<td>Research Question</td>
</tr>
<tr>
<td>TLA</td>
<td>Third Language Acquisition</td>
</tr>
<tr>
<td>SES</td>
<td>Socio-economic Status</td>
</tr>
<tr>
<td>SHS</td>
<td>Spanish Heritage Speaker</td>
</tr>
</tbody>
</table>

SUMMARY

Previous work has identified lexical narrowing and contact-induced lexical change as characteristic of heritage and diaspora Spanish. In order to nuance such attributions, this study performs quantitative and qualitative comparative analyses of the lexical profiles of two generations of Spanish-Swedish-(English) adult bi-/trilinguals residing in Stockholm, Sweden. The data are drawn from timed word association (i.e., lexical availability) tasks that, via prompts, elicited speakers’ vocabulary relating to 20 specific semantic domains. The role of trilingualism in lexical knowledge and borrowing is examined, as well as additional socio-experiential factors, including dialect contact, gender, parental education, exile background, heritage language instruction, and global language proficiency.

The results suggest that diaspora Spanish speakers of the second generation have robust lexical knowledge across a wide range of semantic domains (spanning the domestic and extra-domestic). However, the organization of their mental lexicons (e.g., relationships among words) appears to be less complex than that of first generation speakers. The presence of Swedish and English origin contact lexicon in the speakers' productions was minimal, although for both generations Swedish exerted greater influence than English on the heritage language. The data also suggest that early L3 acquisition of English by the second generation does not threaten heritage language maintenance; on the contrary, stronger L3 English was correlated with greater proficiency in the heritage language. With regard to dialect contact in the form of recurrent travel to Spain, effects of this were observed in the notable presence among the participant productions of certain high-frequency lexical items belonging exclusively to Peninsular dialects. Additional qualitative analyses of the speakers' mental lexicons revealed how first generation conceptualizations of certain semantic fields (e.g., professions; modes of transportation) differed
in striking ways from those of the second generation, likely resulting from divergent socio-cultural upbringings.

This study advances the field's understanding of the lexical knowledge and use of Spanish heritage speakers, challenging its common depiction as "homebound" and "highly restricted", and further explores some of the community-specific factors conditioning language maintenance and change.
1. INTRODUCTION

1.1 Introduction

As noted by Klee (2011, p. 355), the political and economic forces of the last several decades are responsible for a remarkable increase in global migration. According to the United Nations, in 2013 the number of international migrants worldwide had grown to 231.5 million, representing 3.2% of the world’s population (International Migration Wallchart 2013). Latin Americans have been among those significantly affected by such pushes and pulls, evidenced by the expanding presence of Spanish-speaking minorities across the globe during this period, particularly in North America, Western Europe, Australia, and New Zealand (Potowski & Rothman 2011). Having abandoned their homelands due to economic, political, and social difficulties, they face a new set of pressures in the form of cultural and linguistic assimilative demands.

Considering the vital role that language plays in daily transactions and relationships, it is one of the areas most quickly and extensively altered following migration. The language contact situations in Latin American migrant communities have been shown to give rise over time to new varieties of Spanish that reflect a coexistence of languages (Lipski 2008) and dialects (Zentella 1990; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012; Potowski 2008, 2011) while also threatening the maintenance of the heritage language (HL)\(^1\) for younger generations\(^2\). It is well documented that the minority status of Spanish creates a context of reduced input and production for minority language youth (e.g., Silva-Corvalán 1994; Potowski

---

1 A related term used throughout this document is that of Spanish as a heritage language (SHL).
2 Researchers of Spanish in the US have generally followed the generational breakdown established by Silva-Corvalán (1994): First-generation: individuals who arrived after the age of 12; 1.5 generation: individuals who arrived between the ages of 6 and 12; Second-generation: individuals who were born in the US or who arrived before the age of 6; Third-generation: individuals whose parents were born in the US or who arrived before the age of 6. I will use these same criteria to describe the Swedish context of Chilean migration.
2004a) who, although generally orally proficient, readily become dominant in the majority language. The acquisition of the societal language at the expense of the minority language has been shown to impede successful inter-generational language transmission and to generally result in a complete shift to the majority language by the third generation (Bills, Hudson, Hernández Chávez 2000; Valdés 1988).

Though underrepresented in the Hispanic sociolinguistics literature, many transplanted Latin American-origin communities are largely composed of “reluctant migrants” (Poyatos Matas & CuatroNochez 2010), a term used to describe exiles and their posterity. According to Becker (2013, p. 3), “Exiles form a little-investigated immigrant group whose virtual absence in the HLD [heritage language development] literature has gone largely unexplained.” Because they tend to differ in areas such as social class and transnational practices, Becker contends that exile communities have the ability to offer a line of research that “promises to challenge and expand upon what we already know about the processes of heritage language and identity development in other migrant groups.” This is one gap addressed in the present study, which offers a cross-generational analysis of the Chilean diaspora in Stockholm, Sweden, a trilingual HL community with a large refugee presence.

Heritage speakers (HS) are understood as those raised in a home or environment where a non-majority language is spoken and who achieve some degree of bilingualism in the heritage and majority language (Valdés 2000). Because HS can be distinguished from non-HS bilinguals who are strongly dominant in one language – that is, individuals who remain native-like in their L1 but acquire some proficiency in an L2 – it is argued that they provide a missing link between these fully “competent” L1 learners, highly “balanced” bilinguals, and L2 learners (Polinsky 2008). As such, acquiring a complete understanding of the nature of HL acquisition becomes
crucial for the theoretical understanding of language acquisition in general. Furthermore, research into the Spanish of HS is valued for its ability to inform instructional practices for this growing, heterogeneous student population (e.g., Belpoliti & Plascencia-Vela 2013; Lynch 2003; Valdés 1997, 2005; Kagan & Dillon 2008; Fairclough 2011; Valdés & Geoffrion-Vinci 1998).

It is important to note that while attrition (erosion) can affect both adult migrants and their posterity, interrupted or incomplete acquisition (defined as simplified and underdeveloped linguistic features) of the minority language is typically only characteristic of the latter (Benmamoun, Montrul, & Polinsky 2010; Lynch 2003; Montrul 2008; Silva-Corvalán 2003). That is, truncated acquisition and the resulting areas of grammatical simplification and other linguistic divergence are a direct consequence of limited communicative opportunities in the minority language, and in varying degrees they characterize the language experience of all HS. The minority language, including HS varieties, can also undergo significant qualitative changes largely unrelated to attrition or truncated acquisition. These emerge as a result of cultural, conceptual, and other differences between the receiving and sending societies.

HS are worthy of study from many linguistic angles, but they may exhibit particularly rich lexical phenomena given that the effects of language contact and bilingualism seem to be readily and extensively seen in this area (Köpke 2002; Thomason & Kaufman 1988; Weinreich 1953). This includes cases of loanwords, loanblends, calques, and semantic extensions (e.g., Fairclough, forthcoming; Mendieta 1999; Otheguy & García 1993; Poplack, Sankoff & Miller 1988; Smead 1998; Varra 2013), slowed and less accurate lexical access and retrieval (Hulsen 2000; Kohnert & Bates 2002; Montrul & Foote 2014; Polinsky & Kagan 2007; Schmid & Köpke 2009), and a decrease in lexical richness (Fairclough 2005; Schmid 2007). Although there exists a sizeable body of work dedicated to the contact lexicon of Spanish speakers in the U.S.,
vocabulary research specifically within the field of Heritage Language Studies has been largely overlooked in favor of work focusing on the structural characteristics of HS systems (Fairclough, forthcoming). Some of the possible reasons for this disparity are discussed in the next section, as well as in Chapter 2. Relying on the lexical availability methodological framework, this study provides a multi-dimensional, comparative analysis of the lexical productions of two generations of minority Spanish speakers in Stockholm.

1.2 Statement of the Problem

1.2.1 Heritage language acquisition and the lexicon

The growing field of HLS has thus far primarily concerned itself with mapping morphosyntactic structural competence (Benmamoun et al. 2010). As research advances in the understanding of HS grammars, however, there continues to be “less information on the status of the lexicon” (Belpoliti & Plascencia-Vela 2013, p. 67), a gap also noted by Montrul (2010, 2012, 2013) and Fairclough (2011). Although words are often considered the “basic building blocks of language” (Read 2000, p. 1), vocabulary knowledge is a complex construct (Schmitt 2010). Both in terms of lexical breadth (vocabulary size) and depth (knowledge of words beyond simple glosses, including their relationships with other words), we know little about how HS lexical inventories fare under less than optimal acquisition conditions. Such conditions include reduced input and output in the HL and an absence of formal instruction. The over-conflation of language contact phenomena with HS lexical deficiencies (Otheguy & Stern 2011) has seemingly served as a distraction from other aspects of lexical knowledge more germane to proficiency.

The lack of scholarly attention given to HS lexical knowledge may be due to several additional factors. First, there are methodological challenges to obtaining sufficiently
representative quantitative lexical measures, and there continues to be a shortage of HS corpora\(^3\) (Poplack & Dion 2012). There may also be a tendency to ignore gaps in lexical knowledge or to regard them as less vital to language proficiency because the lexicon is a “network of items that are far less densely connected and interdependent” than other closed-class systems and can “tolerate a certain amount of change, loss or interference” without serious structural ramifications (Schmid & Köpke 2009, p. 211-212). Along the same lines, Montrul and Foote (2014) note that the lexicon tends to be ignored in acquisition or attrition research because of its reduced sensitivity to the critical period. Overlooking the lexicon in favor of grammatical competence in language acquisition research may also be a lingering product of the traditional top-down grammar-focused approach to teaching in which “vocabulary development was seen as some kind of secondary or auxiliary activity” (Eyckmans 2004, p. 21).

Nevertheless, considering the substantial research (albeit mostly dealing with English) showing how the lexicon feeds grammatical acquisition and development, both in first- and second-language learning (e.g., Clark 1995; Ellis 1997; Nation 1990, 2001; Pérez-Leroux 2011), as well as emerging correlations between aspects of lexical knowledge and areas of HS grammatical knowledge (Fairclough 2011; Polinsky 1997, 2005), it is clear that the lexical domain merits more scholarly attention. The vital role of the lexical dimension in grammatical mastery, literacy and overall language fluency for different types of learners is discussed in greater detail in Chapter 2.

Despite the inherent complexity of vocabulary knowledge, as well as the factors seeming to set it apart from other types of linguistic knowledge, empirically vague characterizations of SHS lexicon have been put forth in the literature. These describe it as homebound, childlike, and

\(^3\) The Spanish in New York (Otheguy & Zentella 2012), Spanish in Texas (SPINTX) (http://spanishintexas.org), and CHISPA (Potowski & Torres, forthcoming) corpora are examples of rather recent large-scale resources upon which much work has already been based.
thus greatly “restricted” or “compressed” (Blake & Zyzik 2003; Campbell & Rosenthal 2000; Draper & Hicks 2000; Montrul 2012; Polinsky 2008). Montrul (2010, p. 6), for example, depicts SHS vocabulary as mostly limited to words “related to common objects in the home and childhood vocabulary.” In her work on the stylistic range of SHS, Sánchez Muñoz (2010) likewise explains how “as the linguistic repertoire in English expands to include an increasing number of domains, the functional use of the home-based language is restricted to few domains, until it is ultimately limited to the home and family” (p. 337). Although similar reasoning has been put forth as a possible explanation for grammatical simplification (Silva-Corvalán 1996), such assertions regarding the end-state thematic narrowing of SHS lexicon have yet to be sufficiently supported or nuanced by empirical data.

There are at least three somewhat interrelated reasons why the previous descriptions of SHS lexicon may be problematic; First, the domains under the so-called umbrella of “home” and “family” are not well defined in the literature, nor is it clear how they qualify as such. A wide variety of topics can potentially cross the home and family domain, including many that might not necessarily be considered “restricted” or “routine”, undoubtedly vague categories to begin with. Consider, for example, the well-documented language brokering role filled by many children of immigrants in which they are relied on to translate and interpret rather sophisticated discourse (Morales & Hanson 2005), the various forms of Spanish language media openly consumed in the household, or the deeply political discussions regularly taking place in refugee families and communities (Becker 2013). Thus, it appears our understanding of SHS lexical knowledge in terms of domain narrowing requires far more explicit reasoning and evidence.

Secondly, the vast differences in the socioeconomic and migratory histories, language use patterns, and thus proficiencies among speakers (Carreira 2004; González-Pino & Pino 2005;
Potowski & Rothman (2011) render an across-the-board characterization of SHS lexicon somewhat unreliable. Such is the case with many areas of SHS linguistic systems, but the open nature of the lexicon may mean that the range of knowledge across speakers and communities is especially broad. Consequently, in charting the continuum of SHS lexical knowledge, it is important for analyses to consider the diverse and locally pertinent socio-experiential factors that condition language maintenance and development.

Lastly, because lexical acquisition has the ability to freely continue throughout an individual’s lifetime, unlike other language areas that are more sensitive to the critical period (Bayard 1989; Montrul & Foote 2014; Nagy 2011), extra care must be taken to consider how the notions of incomplete acquisition and attrition uniquely intersect with the lexicon. Read (2000, p. 1) describes how continual vocabulary development in adult life happens “in response to new experiences, inventions, concepts, social trends and opportunities for learning.” That vocabulary knowledge “rests heavily on experience and use” explains why it can be largely “spared from the critical period” yet also be highly susceptible to language loss⁴ (Montrul 2008; Montrul & Foote 2014, p. 7). A related perspective is provided by Bolger and Zapata (2011, p. 5), who propose that the allocation of lexical knowledge to declarative memory, as opposed to procedural memory like grammatical knowledge, means that HS lexicon is “more malleable in adulthood and, therefore, more susceptible to the relative frequencies of its overall membership” (that is, to competing L2 vocabulary). A nuanced treatment of SHS lexicon, then, ought to consider how both its propensity and resistance to attrition/underdevelopment might be reflected in speakers’ knowledge. A by-domain analysis of vocabulary knowledge is one way of examining the concomitant effects of these oppositional processes across the lexicon.

---

⁴ This can include attrition both in the form of complete loss or in problems with lexical access.
While not rejecting the realities of truncated development and attrition, this study problematizes the assumption that SHS lexicon is generally restricted to highly personal topics and functions. It does so in the following manner: (1) by testing knowledge across multiple, identifiable semantic domains, (2) considering potentially conditioning socio-experiential factors (e.g.; exile background; parental education; birth order; L3 English proficiency), and (3) investigating the context of a specific community of speakers.

Despite the domain-centered nature of the aforementioned depictions of SHS lexicon (e.g., restricted to the home, childhood topics, etc.), research to date has relied mostly on non-contextualized lexical decision tasks (e.g., Fairclough 2011; Montrul & Foote 2014). In such tasks, words are selected based on frequency levels and presented discretely as a means of determining lexical proficiency. Though a useful tool in measuring general lexical richness, as well as in giving “a guide to the outer limits of vocabulary knowledge” (Cameron 2002), the contribution of word frequency to understanding thematic or context-based knowledge is limited. That is, lexical tests constructed around corpus frequency measures and administered in the absence of context are not built to assess a speaker’s ability to manage linguistic tasks relevant to a particular topic or domain. Nor are they typically designed to provide insight into speakers’ semantic networks (i.e., the relationships among words in a speaker’s mental lexicon) or patterns of lexical variation. They are, however, intended to gauge a speaker’s ability to successfully manage particular genres, such as informal conversation or (non)fiction writing (Davies 2005; Fairclough 2011, forthcoming).

The lack of uniformity across corpora is another potential limitation of measures based on word frequency, such as lexical decision tasks. While derived from quantified data, frequency indexes have been found to be somewhat subjective due to variation in the genres and the
amount of spoken versus written language composing different corpora (Schmitt 2010). In addition, both Gernsbacher (1984) and Nation and Webb (2011) report that a word’s frequency often does not correlate with its level of familiarity for different speakers. Consequently, frequency rates may be misleading in that a word could be of overall low frequency in the corpus counts for a given language, yet in reality be a well known and high frequency (even colloquial) word in everyday speech (Hernández Muñoz, Izura, & Ellis 2006, p. 231; Ferreira & Echeverría 2014). Such corpus subjectivity can result in inflated proficiency assessments or, alternatively, if certain semantic domains are underrepresented in the corpus, results that mask speakers’ knowledge. Germane to this point is the widely accepted notion that bilinguals, even those considered “balanced” in a very general sense, will display differences in dominance by domain (Baker 2011). Thus, complementing frequency-based work with thematically grounded measures can contribute to the creation of a faithful and representative picture of HS lexicon.

In studying SHL development and socialization, Becker (2013) and Gibbons and Ramirez (2004) suggest a shift from the focus on the mere language of use to an examination of language use by topic. This is because not only do topics of interaction themselves vary across and within SHL communities, so does the sophistication of such topics. Consequently, beyond merely providing the opportunity for interaction, the topics of discussion themselves have the ability to facilitate HL developmental processes (Becker 2013, p. 119). For an especially relevant example, both Becker’s (2013) and Gibbons and Ramirez’s (2004) work found that politics was a frequent topic of conversation in Chilean-Canadian and Chilean-Australian households, with members of the second generation utilizing Spanish in political activism and articulately participating in political discourse. This topic, they argue, lends itself to skills in argumentation and other sophisticated discourse, thereby proving particularly beneficial to language development.
Because domains vary in complexity (some involving more technical, abstract, or dense words/ideas than others), an analysis of lexical production by topic can support the endeavor of nuancing the field’s understanding of SHS lexical knowledge. A focus on the contextual lexical knowledge of SHS therefore represents a consideration of current theorizing in HL education, which proposes a shift from a ‘competencies’ perspective to a ‘capabilities’ approach. According to Martínez (2015), it is the context dependent capabilities of HS (i.e., what they are able to do and what they are "able to be" through the HL within their communities), rather than their context independent competence, that should guide the HL teaching profession. Martínez provides as an example how specialized training in medical translation and interpretation for heritage learners can allow them to serve as a bridge between limited English proficiency members of their minority group and healthcare services.

A final acquisition-related aspect of the present study is its focus on lexical production. While the lexical proficiency of HS has been mostly examined in terms of receptive abilities (Fairclough 2011, 2013a; Montrul & Foote 2014; Velásquez 2015), there is evidence that production is more readily and extensively affected by the processes of attrition and truncated acquisition and thus better able to distinguish speaker abilities (Hulsen 2000). Further detail on this difference will be provided in Chapter 2.

1.2.2 Qualitative lexical change

While quantitative analyses of vocabulary size offer important views of the linguistic systems of HS, the idea that the lexicon can undergo diachronic cross-generational qualitative change also merits examination for this speaker group. On the one hand, identifying particular qualitative shifts in the lexicon can support instruction by serving as a guide for the design of materials and curricula. For example, Sánchez Muñoz’s (2010) work suggests that the lexicon of SHS differs
qualitatively from that of Spanish-dominant or monolingual speakers in that it tends to be more monostylistic, i.e., composed of a proportionately large number of familiar, non-technical words. In this sense, qualitative change is tied to a change in lexical complexity. Indeed, the notion that HS systems exhibit stylistic limitations explains why one of the primary objectives in HL education has been to expand their “bilingual range” (Valdés 1997). Vocabulary represents a large part of this expansion, thus warranting special attention in SHL curricular development. If research can help identify particularly strong thematic areas and associated skills (such as argumentation) in HS lexical inventories, educators can work to strategically bridge these with other less familiar domains.

On the other hand, lexical change in the context of multilingualism and multiculturalism is also highly relevant to theories in sociolinguistics and language contact research. Given that not all cross-generational lexical change in a minority language context is attributed to proficiency gaps, it is also of interest from a purely variationist standpoint (Nagy 2011; Varra 2013). The types of lexical variation examined in this study are briefly described ahead.

The qualitative portion of the present study includes, first, an overview of the degree of overlap between generational groups in the lexical items produced in response to six thematic prompts. This includes a discussion of the observed tendencies reflected in the parallel and divergent productions. The six domains analyzed include: *El cuerpo humano* (The Human Body); *Los muebles* (Furniture); *Medios de transporte* (Modes of transportation); *Profesiones y oficios* (Professions and Trades); *La economía* (The Economy); *La política* (Politics). In addition, the participant productions for these six categories are examined to identify (1) the amount and types of lexical transfer from Swedish, English, and Peninsular Spanish; (2) how the most prominent words in the participants’ lists compare between groups, and the socio-cultural
relevance of this kind of variation. The paragraphs ahead briefly describe these aspect; further details are supplied in Chapter 2.

Transfer has been widely observed in the lexicon of diaspora Spanish. As mentioned previously, cross-linguistic influence (i.e., transfer) in the way of loanwords, loanblends, calques, and semantic extensions (e.g., Mendieta 1999; Otheguy & García 1993; Poplack, Sankoff & Miller 1988; Sánchez Muñoz 2010; Smead 1998; Varra 2013) has been shown to readily appear in the Spanish of speakers in situations of societal contact with English. While higher levels of bilingualism have been found to correlate positively with rates of lexical transfer, research has shown that Spanish-dominant (i.e., first generation) speakers also rely on lexical transfer in response to social and expressive needs that emerge in the receiving society (Otheguy & García 1993; Poplack, Sankoff & Miller 1988). Set in the trilingual context of Stockholm’s Chilean community, this study further explores the issue of lexical transfer by asking, first, how transfer varies intra- and inter-generationally when semantic domain is taken into account, and second, how much transfer can be attributed to Swedish (the majority language) versus English (an early L3 for the second generation with increasing societal presence and prestige).

Language contact research has drawn distinctions between core and cultural borrowings (Myers-Scotton 2002), with data showing that initial stages of transfer reflect more extensively the latter, i.e., phenomena related to the customs, culture, and places of the host society (Gamboa González 2003; Otheguy & García 1993). Alternatively, core or basic vocabulary tends to not be affected until advanced stages of language contact (Hock & Joseph 2002; Nagy 2011). However, this pattern has mostly been observed discretely at the word level. The present study looks at borrowing as a function of semantic domain, and analyzes how topics (Moreno-Fernández 2007; Thomason & Kaufman 1988), ranging from the basic/concrete (e.g., el cuerpo humano ‘the
human body’) to the more technical/abstract (e.g., la política ‘politics’), vary in terms of receptivity to foreign-origin lexicon. As our understanding remains limited with respect to the actual density of contact lexicon in minority Spanish (Othe guy & Stern 2011), the cross-domain lexical inventory on which this study is based offers both situated (topic specific) and cumulative (corpus-wide) perspectives on the ratio of contact vs. non-contact lexicon.

The influence of L3 English in the borrowing equation is of special interest. Our knowledge is limited in terms of what lexical transfer in minority Spanish looks like in the context of two societally prestigious languages (de Bot & Gorter 2005). That is, how does trilingualism in which two of the languages, Swedish and English, enjoy societal status and substantial public presence shape the lexical repertoire of the less prestigious heritage/minority language (Spanish)? With respect to the SHS population of Stockholm, many questions arise: How might proficiency in English influence their vocabulary in Spanish, both on a grand scale (i.e., the statistical effect of degree of English proficiency on vocabulary size in the HL) as well as in terms of the presence of Anglicisms in different thematic areas (e.g., clothing, food, entertainment, health, etc.)? What do such patterns reveal concerning the expanding role of English in Swedish society? Relatedly, from a comparative perspective, how might the types of English borrowings in this speaker population compare with those observed in the productions of U.S. SHS on a similar task (Moreno-Fernández 2007)? Despite low levels of trilingualism among G1 Chilean Swedes, does the salience of English in Swedish society nonetheless impact their lexical repertoires?

Under the umbrella of qualitative change, the influence of Peninsular Spanish on the variety of Chilean Swedes is also examined. While there is a small body of research on the effects of dialect contact on SHL development in the U.S. (e.g., Potowski 2008; Ramos-Pellicia
2004; Raymond 2012), including through formal instruction (Fairclough 2005), we know little of the dialectal influences on minority Spanish brought about by extended travel to other-dialect Spanish-speaking regions and countries. Whereas travel to Chile is costly, emotionally taxing (given their exile background), and time-consuming, Spain is an accessible, low-cost, language-compatible destination for many vacationing Chilean Swedes. Given the Swedish culture of frequent travel and generous vacation time, Spain, particularly its southern regions, is a place many visit on a regular basis. In this study I test whether regular contact with Peninsular Spanish varieties is having an appreciable effect on the Spanish lexicon of the Chilean minority in Sweden. Research has shown that in second dialect acquisition, the lexicon is first to be acquired (Siegel 2010). Also, in general, the younger the speaker, the greater propensity there is for change as a result of persistent dialect contact. Thus, the study will take into account generational differences in this regard.

Another area of the lexicon that may undergo inter-generational qualitative change is that of its mental organization as a reflection of category conceptualization. Aitchison (1994) discusses a “mental model” of the lexicon that results from several factors, such as observation and experience, cultural influence, memory, and imagination. In general, categories play a large role in human cognition. As a result of repeated experiences, categories of natural objects (e.g., animals) or concepts (e.g., colors) are formed and become organized around prototypes (Rosch 1973, 1978). Sharifian (2008) describes how “human conceptual faculties derive from various sources of experience, including bodily and environmental, which enable new experiences to be made sense of and organized” (p. 9). In other words, the complexity and systematicity of our knowledge is aided by the formation of units of organization in our minds, such as categories,
which, according to Robinson (1997, p. 263) reflect “regularities in an organism’s perception of and interaction with its environment”.

People of different cultural and linguistic backgrounds can have strikingly diverse construals of the self, of others, of nature, of social structures, etc. These varying conceptions can influence, and even determine, the very nature of individual experience, including cognition (Lakoff 1999). Theories of how aspects of one’s native language influence thinking and conceptualizations of the world are based on a similar premise (Boroditsky 2001; Kurinski & Sera 2011; Nicoladis & Foursha-Stevenson 2012; Sapir-Whorf Hypothesis). Drawing on Rosch’s (1978) theories of categorization and prototype levels, scholars have explored cross-linguistic conceptual discontinuities in different thematic areas: e.g., emotions (Lakoff & Kovecses 1987; Kashkooli 2014; Shaver et al. 1987; Wierzbicka 1999), colors (Moore, Romney & Hsia 2000), and others, including the concept of a lie (Holland & Quinn 1987). Other work has investigated how cultural cognition compares across speakers of different varieties of the same language (Carcedo González 2000; Jarvis & Pavlenko 2008). Polzenhagen and Xia (2014, p. 19) explain how embodied and encultured categorization differ in these processes:

On the one hand, the shared nature of our body assures that both the basic principles we abide by in categorizing the world and the fundamental structure of our categories are universal. On the other hand, the cultural experience and knowledge we gain from our situatedness in a particular culture inevitably shapes our thoughts, making the conceptual categories not only embodied but also ‘encultured’. Under a crosscultural perspective, prototype theory predicts not only that translation-equivalent categories from two cultures exhibit a centre-periphery structure, but also that the prototype and prototypicality gradation among the category members are culturally situated.
Given that the organization of the mental lexicon is bound up with early experiences and cultural-environmental influences, it is conceivable that, in a minority language context, speakers’ conceptions of semantic categories vary across generations. In the Chilean Swedish community, the socialization and environmental upbringing of G1 adult migrants in Chile differs on many levels from what their G2 children have experienced in the Swedish context. Therefore, there appears to be a great propensity for the categorical conceptualizations of the Swedish-raised generation to diverge from those of their parents. The present study shows how the prominence (i.e., prototypicality) of particular items/concepts associated with a given semantic category can vary considerably between generations. This type of socio-culturally shaped cognitive shift in a diaspora speech community represents an aspect of language change perhaps far less visible than lexical borrowing, yet one that is also relevant to understanding lexical variation within the context of contact. The inclusion of a variety of categories in the experiment served to explore whether a category’s level of abstraction favors or discourages this type of variation. In Chapter 2, theories of cultural conceptualization and culture-specific prototypes will be discussed in greater detail, including concrete examples.

1.2.3 Goals of the study

To examine the quantitative and qualitative nature of SHS lexicon, I perform a multi-domain comparative analysis of the productive lexicon of two generations of Chilean-origin speakers in Stockholm, Sweden. To provide further perspective, portions of the data are also measured against comparable findings for homeland monolingual speakers (reported in Valencia & Echeverría 1999). A detailed profile of the Chilean-Swedish community is presented in Chapter 2. These speakers represent both an understudied Spanish-speaking diaspora and language
contact pair (Spanish-Swedish) in the literature. In addition, the presence of English as an L3 creates a trilingual factor largely unexplored in research on diaspora/heritage Spanish.

The data were obtained through the employment of tests of lexical availability, which provide a type of lexical inventory for the speech community in question. This methodology consists of timed word association tasks that elicit vocabulary lists based on thematic prompts corresponding to various domains of experience. Quantitative and qualitative aspects of the words produced in these tasks combine to detail (and facilitate comparisons between) the productive vocabularies representative of each speaker and speaker group. The prompts presented to participants were selected with the aim to cover a broad thematic range, from basic topics common to the concrete spheres of home and family (muebles ‘furniture’; ropa ‘clothing’), to more abstract topics tied to broader society (la política ‘politics’; problemas sociales ‘social problems’)\(^5\). Further methodological details are provided in Chapter 3.

Moreno-Fernández (2007, 41) describes available lexicon as “(1) the sum of words that speakers have in their mental systems and (2) whose use is conditioned by a particular topic”. Similarly, Hernández Muñoz (2009) defines it as “el conjunto de palabras que son más fácil y rápidamente producidas por los hablantes en torno a un tema determinado” (p. 224). Tests of lexical availability are primarily used to determine: (1) the number of words speakers produce as related to a given semantic domain; (2) the types of words produced; and (3) the prevalence that some words have over others in production. The first and second are relevant for questions concerning proficiency, which is a primary focus of this study. The second and third have dialectological importance, and the third provides perspective on the socio-cognitive structure of the mental lexicon – lenses that my analysis will use to shed light on language variation and change in the Chilean-Swedish community.

\(^5\) In Chapter 3 I provide a complete list of the 21 domain prompts used.
I will describe the concept and methodology of lexical availability, as well as provide terminological definitions (*word, lexical item*, etc.) in Chapters 2 and 3, but one clarification is in order at this point: In this work, the terms lexical knowledge, lexical proficiency, and lexical availability are all used interchangeably in particular reference to performance on the word association tasks used in lexical availability research. The following are the primary questions guiding my study:

1. What are the quantitative characteristics of SHS lexicon, as measured in a lexical availability task?
   a. Is there a significant numerical difference between SHS (G2) and first generation (G1) speakers in the words produced for each of 21 semantic domains? What role does domain complexity play?
   b. What is the relationship between SHS lexical knowledge (number of words produced) and the following social variables: gender; birth order; parental education; and exile background?
   c. What is the relationship between SHS lexical knowledge (number of words produced) and the following language variables: amount of HL instruction; performance on a general Spanish proficiency measure; performance on an English proficiency measure; self-ratings of Spanish and English proficiency; media exposure to these languages; and Spanish language use?

2. What are the qualitative lexical characteristics of the participant productions in six semantic domains? Specifically, how do the following reflect inter-generational change?
   a. Compatibility rates and characteristics of divergent lexicon
b. The density and types of Swedish and English borrowings

c. The density and types of Peninsular Spanish regionalisms

d. Category conceptualizations

Drawing on diverse methods/perspectives from the fields of language acquisition, sociolinguistics, language contact, dialectology, language change, lexico-statistics, and psychology, this analysis advances understanding of the nature of SHS lexical acquisition and use, the social and experiential factors conditioning vocabulary knowledge, and the types of lexical change occurring in the Chilean Swedish community. An examination of the data through these various analytical frameworks exemplifies what Pires and Rothman (2009), in reference to HL acquisition, describe as “an ideal locus for the integration of methodologies from different areas of linguistic inquiry” (p. 36). In addition to its contribution to addressing what is still in many respects the “uncharted territory” (Polinsky 2007) of HS systems, this project sheds light on an understudied Hispanophone diaspora, language contact setting, and community type (exile).
2. LITERATURE REVIEW

2.1 Heritage Speakers

The label *heritage language* lacks a universally accepted definition. Fishman (2001) categorized HLs in the United States as colonial, indigenous, or immigrant languages other than English. Montrul (2008) notes that a common thread among HLs is that they are ascribed little prestige, because they are “languages associated with the ethno-cultural heritage of particular minority populations, which for many historical and political reasons have been devalued” (p.1). In other words, given the strong monolingual linguistic culture of the U.S., HLs often hold little clout simply because they are not English (Potowski 2010).

In the U.S., the lay population has often employed the basic term *native* to include individuals who, although raised and educated in the U.S., have developed proficiency in a non-English language in their youth, usually through input in familial contexts. Some researchers have considered this term unsuitable because the language of these speakers can differ in important ways from that of monolingual/monolingually raised speakers (Fairclough 2005; Peale 1991), or because the word *heritage* invokes the past (García 2005). Other terms used to describe these speakers have included *bilinguals* and *heritage speakers* (HS). Since around 1996, following the Canadian tradition, there has been an almost categorical reference in the literature to this speaker group using the latter term, i.e., *heritage language speakers or learners*6 (Beaudrie, Ducar, & Potowski 2014). Alternatively, in the Swedish context, the terms of common usage are *mother tongue speaker*, or simply *bilingual*. In order to remain constant with the bulk of current literature on bilinguals fitting this profile, in this study I employ the term *(Spanish) heritage speaker* (S)HS to denote bilinguals of the second generation (G2), defined as

---

6 Hence the organizational name *National Heritage Language Resource Center* and its quarterly publication, *Heritage Language Journal*. 

those born in Sweden or brought to Sweden before age seven. First generation speakers (G1), defined as migrating to Sweden after the age of 12, are variably referenced either by their generational label (G1) or as Spanish-dominant or the parent generation.

Regarding the HS designation, a widely cited definition is that of Valdés (2000), who defines an HS as an individual “who is raised in a home where a non-English language is spoken, who speaks or merely understands the heritage language, and who is to some degree bilingual in English and the heritage language” (p. 1). According to this definition, the commonality among HS is that they possess at least some knowledge and functional ability in their HL gained from their home and/or community environment. Other definitions are much broader. For example, in her learner typology, Carreira (2004) includes learners who have very limited skills in the HL yet strong family connections to it. Similarly, Van Deusen-Scholl (2003) calls these students “learners with a heritage motivation”. Au et al. (2002) employ the term overhearers to refer to HS who frequently heard the language as children but did not develop fluency; interestingly, their research has found that such “overhearing” in childhood has significant advantages in later learning of the language. Because the present study deals primarily with advanced linguistic knowledge and production, all reference to HS will be in accordance with Valdés’ (2000) narrow proficiency-based definition, substituting Swedish for English.

From a purely linguistic standpoint, HS are often categorically grouped due to shared aspects of their grammars, which researchers argue show signs of three phenomena: incomplete acquisition, due to interrupted learning and early dominance in the L2; attrition, due to lack of use and input; and contact features, resulting from high levels of bilingualism and intense

---

7 Generational criteria follow the typology presented in Silva-Corvalán (1994).
8 This definition can be extended to include speakers of non-majority languages in any society.
contact with a majority language. These phenomena are often difficult, if not impossible, to tease apart reliably. Furthermore, some socially oriented scholars dispute the existence of incomplete acquisition. For example, Green and Morgan (2005) contend that an individual’s linguistic system is never truly “complete.” Similarly, Otheguy and Zentella (2012) insist that, given our limited theoretical understanding of what would make a grammar “complete” (a term they argue lacks a definition and even proof of existence), as well as the pervasive fact that language is constantly in a state of flux in all natural speech communities, it would be more accurate to describe the grammatical knowledge of HS as simply different from that of monolinguals.

In spite of these problems, psycholinguists defend the position that complete acquisition is possible in the sense that “native speaker competence and performance are typically the result of normal first language acquisition in a predominantly monolingual environment, with optimal and continuous exposure to the language” (Benmamoun et al. 2010, 1). In this view, HS grammars are regarded as incomplete, due to reduced input conditions brought about by early bilingualism and socialization/education in the majority language. Under these circumstances, developmental stages typical in childhood can persist into adulthood, preventing age-appropriate proficiency and an ultimate attainment of the grammar. HS nevertheless comprise an indisputably heterogeneous population with proficiencies spanning the merely receptive to the fully productive (Bermel & Kagan 2000; Beaudrie, Ducar, & Potowski 2014; Carreira 2004).

While holding that SHS competence is often underestimated, this study relies on the theoretical notion of incomplete acquisition but also uses notions of change in the cross-generational examination of proficiency-related aspects of the lexicon of Chilean-Swedes. A cross-generational analysis allows for the parental generation to act as the baseline for

---

9 Examples of this reasoning include the loss of the who-whom distinction in English, or of the future subjunctive in Spanish, which do not constitute evidence of “less complete” linguistic systems than those of the past.
comparison. Although imperfect, this is considered the most appropriate baseline given its years of contact with the majority language as well as its being the variety HS have been exposed to as their primary source of input. To use the standard available to monolingual/monolingually-raised speakers would be wholly unrealistic (Benmamoun et al. 2010, 10; Polinsky & Kagan 2007).

2.2 The Chilean Diaspora and the Chilean Minority in Sweden

Before proceeding to a review of research on the lexical characteristics of SHL (Section 2.4), I will provide a general overview of the Chilean diaspora to contextualize the experiences of my study participants. I will also briefly discuss the linguistic research to date on Spanish-Swedish bilingualism.

As a result of the military coup of 1973 in Chile, which set in motion the 16-year dictatorship of the Chilean military ruled by General Augusto C. Pinochet, approximately 200,000 Chileans were forced into exile. As detailed by Wright and Oñate (2005, p. 57), Pinochet’s regime “used state terror to seize and retain control, systematizing the violation of human rights by employing arbitrary detention, torture, murder, and disappearance against those it deemed enemies.” Such treatment by the dictatorship toward all “real or perceived opponents” led to the Chilean diaspora, and prepared the way for similar rebellion and state terror in neighboring countries (with their own subsequent histories of mass exodus). Wright and Oñate (2005) relate how immediately after the coup, many sought asylum in embassies and received assistance to enter safely into exile. However, the larger part of Chilean exiles departed between 1974 and 1976 as an alternative to a long prison sentence or under increasing pressure and threats from the National Intelligence Directorate (DINA). The mass exile carried negative economic repercussions, resulting in a second exodus in the 1980s and early 1990s, despite the recent restoration of political democracy.
Wright and Oñate (2005) recount how although at the time of the coup economic and political circumstances in the neighboring countries of Peru and Argentina were far from ideal (having similar histories of political unrest), these were the first stops for many exiles. Mexico, Venezuela, Cuba and Costa Rica were especially welcoming. Due to its thriving economy, Brazil also admitted a large number of skilled Chileans. In North America, a significant number also settled in Canada, whereas the United States’ pivotal role in the overthrow correlated with their acceptance of few. Since the 1990s, however, voluntary expatriates in the U.S. have risen to number more than 113,000 (Chilean Ministry of Foreign Affairs and National Institute of Statistics 2005). These more recent migrants have tended to be more socially mobile than the earlier exiles. In Europe, the policies of Sweden and France allowed for the settlement of a significant number of exiles, with the larger urban areas as their preferred places of residence. Today, the number of foreign-born Chileans and their descendants in Sweden exceeds 40,000. Other large Chilean diasporas outside of Latin America are found in Australia, Canada, and Spain. In descending order, Table 2-1 displays the countries with the largest Chilean populations as of 2005.

<table>
<thead>
<tr>
<th>Country</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>429,708</td>
</tr>
<tr>
<td>United States</td>
<td>113,394</td>
</tr>
<tr>
<td>Sweden</td>
<td>42,396</td>
</tr>
<tr>
<td>Canada</td>
<td>37,577</td>
</tr>
<tr>
<td>Australia</td>
<td>33,626</td>
</tr>
<tr>
<td>Brazil</td>
<td>28,371</td>
</tr>
<tr>
<td>Venezuela</td>
<td>27,106</td>
</tr>
<tr>
<td>Spain</td>
<td>23,911</td>
</tr>
<tr>
<td>France</td>
<td>15,782</td>
</tr>
<tr>
<td>Germany</td>
<td>10,280</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>762,151</strong></td>
</tr>
</tbody>
</table>
In 2005 a large-scale survey of Chilean-born and Chilean-origin individuals abroad was published by the Chilean Ministry for Foreign Affairs and Chilean National Institute of Statistics. One finding indicative of the generally established nature of the Chilean diaspora is that of all Chileans living abroad, more than 57% had resided in their current location for at least 20 years. The following are a few significant findings on Chileans in Sweden:

- 55% had adopted Swedish nationality (compared to only 22.5% in the U.S.), 62% of whom desire to recover their Chilean nationality
- 91.4% still had direct family in Chile (parents, siblings, children, grandparents)
- Only 17.5% visited Chile every year or two
- 41.9% had intentions to return to Chile to live (mostly those 40-64 years of age, and slightly more men)
- 32.9% belonged to a Chilean organization (a greater proportion than in any other Chilean diaspora): sports organizations – 28%; culture/art groups – 28%; solidarity groups – 20%; political associations – 13.5% (also the highest among all Chilean diasporas)

The Spanish speakers with whom Chilean-Swedes primarily interact are fellow Chileans, although there are other notable Spanish speaking groups, including Colombians, Peruvians, Bolivians, and Spaniards. Spain represents a popular vacation destination for Chilean Swedes, where they experience further linguistic contact with Peninsular dialects. Section 2.7.2 provides further details on this particular context of dialect contact. The numerical dominance of Chileans within the Latin American population in Sweden beginning around 1975 is shown in Table 2-2,
which displays the growth of the foreign-born Spanish speaking population in Sweden over more than a century. Continued migration from Chile is minimal, however, and post-retirement returnees continue to affect population counts. Currently, Chileans represent 35% of all foreign-born Spanish speakers in Sweden (unfortunately, data including descendants of foreign-born individuals, i.e., Swedish-born children and grandchildren of immigrants, is limited).

**TABLE 2-2**
FOREIGN-BORN SPANISH SPEAKING POPULATION IN SWEDEN BY COUNTRY OF ORIGIN 1900-2014

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>26</td>
<td>83</td>
<td>197</td>
<td>272</td>
<td>632</td>
<td>640</td>
<td>2211</td>
<td>2326</td>
<td>2421</td>
<td>2441</td>
<td>2435</td>
<td>3114</td>
</tr>
<tr>
<td>Bolivia</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>12</td>
<td>327</td>
<td>405</td>
<td>613</td>
<td>1673</td>
<td>1983</td>
<td>2231</td>
<td>2349</td>
<td>4386</td>
</tr>
<tr>
<td>Chile</td>
<td>6</td>
<td>28</td>
<td>30</td>
<td>69</td>
<td>181</td>
<td>688</td>
<td>660</td>
<td>2211</td>
<td>2326</td>
<td>2421</td>
<td>2441</td>
<td>2435</td>
</tr>
<tr>
<td>Colombia</td>
<td>-</td>
<td>-</td>
<td>20</td>
<td>73</td>
<td>259</td>
<td>531</td>
<td>1443</td>
<td>2910</td>
<td>4560</td>
<td>6417</td>
<td>7317</td>
<td>11709</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>15</td>
<td>29</td>
<td>68</td>
<td>105</td>
<td>133</td>
<td>180</td>
<td>207</td>
<td>415</td>
<td></td>
</tr>
<tr>
<td>Cuba</td>
<td>-</td>
<td>-</td>
<td>20</td>
<td>76</td>
<td>110</td>
<td>139</td>
<td>166</td>
<td>874</td>
<td>1489</td>
<td>2653</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dom. Rep.</td>
<td>1</td>
<td>-</td>
<td>4</td>
<td>29</td>
<td>53</td>
<td>73</td>
<td>118</td>
<td>174</td>
<td>225</td>
<td>245</td>
<td>708</td>
<td></td>
</tr>
<tr>
<td>Ecuador</td>
<td>9</td>
<td>-</td>
<td>9</td>
<td>55</td>
<td>221</td>
<td>463</td>
<td>530</td>
<td>647</td>
<td>805</td>
<td>950</td>
<td>2353</td>
<td></td>
</tr>
<tr>
<td>El Salvador</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>121</td>
<td>97</td>
<td>471</td>
<td>1721</td>
<td>2302</td>
<td>2332</td>
<td>2883</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>11</td>
<td>52</td>
<td>93</td>
<td>145</td>
<td>366</td>
<td>495</td>
<td>631</td>
<td>658</td>
<td>810</td>
</tr>
<tr>
<td>Honduras</td>
<td>-</td>
<td>-</td>
<td>9</td>
<td>2</td>
<td>224</td>
<td>42</td>
<td>73</td>
<td>106</td>
<td>167</td>
<td>184</td>
<td>210</td>
<td>369</td>
</tr>
<tr>
<td>Mexico</td>
<td>3</td>
<td>15</td>
<td>45</td>
<td>62</td>
<td>168</td>
<td>265</td>
<td>364</td>
<td>545</td>
<td>696</td>
<td>858</td>
<td>1079</td>
<td>2432</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>33</td>
<td>37</td>
<td>71</td>
<td>99</td>
<td>165</td>
<td>201</td>
<td>201</td>
<td>297</td>
<td></td>
</tr>
<tr>
<td>Panama</td>
<td>-</td>
<td>-</td>
<td>12</td>
<td>6</td>
<td>37</td>
<td>71</td>
<td>99</td>
<td>165</td>
<td>201</td>
<td>201</td>
<td>297</td>
<td></td>
</tr>
<tr>
<td>Paraguay</td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>21</td>
<td>98</td>
<td>125</td>
<td>135</td>
<td>128</td>
<td>130</td>
<td>199</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peru</td>
<td>5</td>
<td>-</td>
<td>19</td>
<td>17</td>
<td>148</td>
<td>305</td>
<td>643</td>
<td>1022</td>
<td>1783</td>
<td>3548</td>
<td>4531</td>
<td>7398</td>
</tr>
<tr>
<td>Uruguay</td>
<td>4</td>
<td>-</td>
<td>13</td>
<td>30</td>
<td>71</td>
<td>477</td>
<td>2101</td>
<td>2624</td>
<td>2430</td>
<td>2265</td>
<td>2227</td>
<td>2172</td>
</tr>
<tr>
<td>Venezuela</td>
<td>-</td>
<td>7</td>
<td>-</td>
<td>16</td>
<td>135</td>
<td>152</td>
<td>223</td>
<td>230</td>
<td>319</td>
<td>381</td>
<td>449</td>
<td>1111</td>
</tr>
<tr>
<td>Subtotal Latin America</td>
<td>54</td>
<td>133</td>
<td>395</td>
<td>613</td>
<td>2538</td>
<td>5037</td>
<td>17052</td>
<td>26672</td>
<td>45929</td>
<td>54696</td>
<td>54055</td>
<td>71999</td>
</tr>
<tr>
<td>Spain</td>
<td>30</td>
<td>64</td>
<td>197</td>
<td>867</td>
<td>3781</td>
<td>3816</td>
<td>4363</td>
<td>4618</td>
<td>4917</td>
<td>4996</td>
<td>5079</td>
<td>9495</td>
</tr>
<tr>
<td>Total</td>
<td>84</td>
<td>197</td>
<td>592</td>
<td>1480</td>
<td>6319</td>
<td>8853</td>
<td>21415</td>
<td>31290</td>
<td>50846</td>
<td>59692</td>
<td>59134</td>
<td>81494</td>
</tr>
</tbody>
</table>

Source: Statistics Sweden, 2014; Note: Spanish-speaking abilities are inferred from the country of origin data.

Today Chilean-Swedes have entered the third generation, and as a group have undergone significant acculturation in a post-exile state (Olsson 2009). Given the amount of time they have been in Sweden and the limited number of new arrivals over the last two decades, one important shift involves their residential patterns. Upward mobility and societal integration have resulted in
less clustered living than in previous decades and a diminished presence in the immigrant-dominated neighborhoods (now largely inhabited by more recent African and Middle-Eastern arrivals). Such movements can carry linguistic consequences, as community has traditionally been known to play an important part in language maintenance (Alba et al. 2002; Fishman 1966; Guardado 2008; Phinney, Romero, Nava, Huang 2001). Alba et al. (2002), for example, found that a Cuban-origin child growing up in Miami is 20 times more likely to become proficient in Spanish than a child living in an area where just 5% of the residents speak the language. Communal factors are also tied to ethnic identity, which in turn correlates importantly with language use and proficiency (see Phinney et al. 2001). While they signal advancement and opportunities on many levels, factors such as these (i.e., stagnant population; residential diffusion) do not favor the HL maintenance and bicultural development of the newer generations.

Olsson (2009) reports on the dispersing effects of a more socially diverse (exiled and non-exiled immigrants) and less romanticized (no longer awaiting repatriation) post-exile phase. During Pinochet's regime, Chileans in Sweden were nearly exclusively political exiles who shared the commitment to Chilean political activism from afar as well as the vision of returning to Chile once Pinochet was removed from power. As time elapsed with little change in the homeland political climate and an inevitable integration into Swedish society, the desire and ability to return waned. With the eventual 1990 resignation of Pinochet’s regime, the decision of many to stay in Sweden ushered in a post-exile phase. This phase has broadened the boundaries of the Chilean-Swedish experience (now less focused around an exile identity), enabling a more accessible participation in the community by the newer generation and more recent non-refugee migrants. This “reorientation of diasporisation” (Olson 2009, p. 670) has meant that, to a certain extent, homeland political activism has been replaced with greater involvement in Swedish social
issues as well as an expanding commodification and consumption of Latin American products (goods, foods, popular culture, folklore, etc.).

A few studies to date have examined the ethnolinguistic identities and language use of Chilean youth, as well as the linguistic consequences of Spanish-Swedish contact. I will now briefly review this body of research and discuss how the SHS context in Sweden compares and contrasts with its U.S. counterpart.

2.2.1 Transmigration and the national, cultural, and linguistic identification of Chilean-Swedes

In relation to the identification processes of Chilean-Swedes, Ganuza and King’s (2005) contribution provides nuanced and insightful perspectives. The authors conducted qualitative interviews to survey the language use, attitudes and identity constructions among bilingual Chilean youth in Stockholm. More specifically, their analysis includes a focus on attitudes toward Spanish, Swedish and Spanish-Swedish code-switching, and investigates the role of language, culture and nationality in the development and maintenance of transmigrant identities.

Ganuza and King (2005, pp. 191-193) describe the participants’ attitudes as seemingly "supported by ideologies of linguistic purity, as well as fears of being viewed as less competent Spanish and Swedish speakers as a result of their code-switching." The participants exhibited the tendency to “assess themselves as not being satisfactorily or sufficiently ‘Swedish’ or ‘Chilean’ and, simultaneously, not speaking good enough Swedish or Spanish.” While the participants’ discourse suggested they were forming identities that were both Swedish (part in-group) and Chilean (part out-group), their descriptions also pointed to a feeling of belonging to two worlds without “feeling themselves to be full-fledged members of either”. Gender was found to impact the degree of affiliation with the majority culture and language, with the boys drawing more attention to their experiences with discrimination than did the girls. The authors report that,
surprisingly, time spent in Sweden did not always correlate positively with “feelings of Swedishness”, particularly in the case of the male participants. Other factors found to contribute to a lack of investment in a Swedish identity for some Chilean migrant youth were an ambiguous legal status, an indefinite future in Sweden and segregated residential patterns.

Lastly, the authors observed that the inability to make regular visits to Chile due to the great distance, finances, or politics contributed to “idealized visions of life in Chile”, a place many participants had never visited, and to idealized notions of “true” Chilean identity (dark, lively, fun-loving) versus sterotypical Swedish characteristics (blond, humorless).

2.2.2 Maintenance, attrition, and Spanish-Swedish contact

For their research on maturational and other constraints in language learning and attrition, researchers at the Centre for Research on Bilingualism at Stockholm University have relied on local Spanish-Swedish bilingual participants. Bylund, Abrahamsson and Hyltenstam (2010) investigated the role of language aptitude in attrition through a series of tests administered to Chilean-origin early Spanish-Swedish bilinguals (SHS) as well as to a control group of monolingual Chileans in Chile. Although all of the target group speakers were considered bilingual, the authors found that an above-average language learning aptitude score correlated with more accurate responses on a grammaticality judgment test. However, speaking to the powerful impact of input, they also reported a significant degree of accuracy for low aptitude scorers that had a large amount of daily L1 input. Thus the conclusion was that “language aptitude has a compensatory function in situations of reduced L1 contact, in that the speaker’s degree of aptitude to a certain extent regulates his/her dependency on L1 contact to achieve and maintain L1 proficiency” (p. 459). On the other hand, another study by Bylund and Ramírez-Galan (2014) reported a lack of aptitude effects on degree of L1 attrition for late Spanish-
Swedish bilinguals (i.e., adult migrants). For this group, the only significant predictor of task performance was linguistic identification. That is, the greater their identification with their L1 (Spanish), the less attrition was found to have occurred.

Other work with this population has looked at instances of contact-induced language change. Bylund (2011), for instance, examined the temporal structuring of events in the plot retellings of early Spanish-Swedish bilinguals. He found distinct cases of “convergence and coexistence of conceptual patterns” from both languages, in line with Pavlenko’s (1998, 2008) work on the language and cognition of bilinguals. Bravo-Cladera (2010), using recorded spontaneous speech of SHS youth in Sweden and monolingual adolescents in Chile, analyzed the impact of Swedish in the use of discourse markers and linguistic feedback items (LFI), or filled pauses, defined as “interaction regulators” that indicate understanding of what is being said (p. 140). She found the bilinguals’ behavior to differ significantly from that of the monolinguals. Specifically, she reported a frequent use of the Swedish LFI$s Mm$ and $Mhm$ as “continuers of interaction” and “acknowledgement of comprehension”, which she attributed to the adolescents’ early acquisition of Swedish.

Parada (2011) examined the adjective placement patterns of adult SHS in Stockholm. Positing the relevance of the Interface Hypothesis for cross-linguistic influence at the syntax-semantic/pragmatic interface of adjective placement, 11 HS and seven Spanish-dominant or monolingual controls completed a series of four tasks designed to elicit their use of pre- and postnominal evaluative adjectives. The results pointed to a simplification of adjective placement rules, chiefly in the diminished reliance on the prenominal position by the HS group when pragmatic conditions called for it. In other words, the participants displayed a regularization of the adjective to the more frequent postnominal slot, as well as less sensitivity to nuances
associated with and expressed through the prenominal adjectival position (e.g., figurativeness, presupposition, sarcasm). Interestingly, a hierarchy was observed in which semantically overt prenominal adjectives (e.g., pobre, viejo, bueno, etc.) were appropriately placed to a greater extent than adjectives requiring preposing for pragmatic or stylistic purposes. While knowledge of Swedish generally did not result in any calquing of adjective-noun word order, the HSs’ movement toward a single-spot rule for adjective placement is taken as evidence for borrowing at the rule level. Parada suggests that attention be given to the role of the participants’ early L3 English in the progression of such processes, given the typological similarity between Swedish and English.

Gamboa González (2003) is an especially relevant study for the present investigation given its focus on contact lexicon. He studied the presence of Swedish lexical loans in the Spanish of 40 Chilean repatriates of varying ages who spent between 12 and 20 years in Sweden before returning to Chile. The group consisted of eight individuals who were either born in Sweden or who had arrived to Sweden before the age of 2, and 32 who had arrived as adults, mostly between the ages of 18 and 26. Of the eight individuals raised in Sweden, two had completed university studies in Sweden while the remaining six were in high school at the time of the family's repatriation to Chile. The data collection took place within ten years of the participants’ return to Chile. Gamboa González obtained data through recorded interviews and field notes, which he transcribed and surveyed for loanwords. He classified the loanwords according to Haugen’s (1972) typology, which included the categories of loanwords (phonologically adapted borrowings), loanblends (morphologically adapted borrowings), and calques.
In his corpus the 40 speakers produced 271 cases of loanwords, 14 loanblends, and one calque. Consistent with other research was the dominance of nouns (62%) among the borrowings, followed by verbs (10%) and adjectives (6%). This pattern is attributed to the well-documented fact that loanwords are often used to describe culturally specific objects or concepts, which are usually nouns. While both generations of speakers produced borrowings, the younger group produced them with greater frequency, although no statistics were reported.

The extent and type of the borrowings produced reflected an intimate knowledge and relationship with the Swedish language and culture, gained through up to 20 years of living in that country. Although the speakers produced loanwords belonging to several different domains (e.g., health, education, parties and celebrations, food and drink, technology, politics, furniture/home, books/publications), Gamboa González noted that speakers resorted to Swedish mainly when there were gaps in the language dealing with phenomena related to Swedish customs, culture and places (e.g., *midsommar* ‘summer solstice celebration’, *smorgasbord* ‘buffet’, *korv* ‘sausage’, *stadsbibliotek* ‘city library’. The need to borrow seemed to apply even when there appeared to be a translation equivalent in the L1, such as with the concepts of *korv* and *smorgasbord*, which the author attributes to a lack of one-to-one correspondence between speakers’ conceptualizations. These findings are in line with other research (e.g., Otheguy & García 1993) in which emphasis is placed on the emergence of lexical gaps due to the contact of cultures and concepts, and the expressive inadequacy speakers would face without the aid of borrowing.

The work outlined in this section reflects the scholarly interest in Spanish-Swedish bilingualism, yet also sheds light on the gaps in our linguistic understanding of this community. One such gap is the lexicon of Chilean diaspora Spanish (as used by non-returnee speakers
remaining in Sweden), including the distribution of contact effects from Swedish, English, and other Spanish dialects. Unaddressed in all of the publications surveyed, the broader role of L3 English in HL maintenance/shift requires examination as well. Considering the interesting connections found thus far between gender, generation, cultural affiliations, and attrition, further exploration of the social correlates of language development and attrition will advance research in fruitful directions. As a variety with a finite existence ahead, we would do well to treat the Spanish spoken by Chilean-Swedes as an aging variety, the study of which carries a certain urgency.

In contrasting the SHS experience in Sweden with that in the United States, there appear to be several distinctive aspects to consider with respect to the former: reluctant migration and exile communities (Becker 2013; Poyatos Matas & CuatroNochez 2011); distance from the homeland, both physically and figuratively (e.g., minimal imported or homegrown Spanish language media); negligible present-day migration from the homeland; the “minority among minorities” status of Hispanics and the Spanish language; and naturalization laws (no citizen-by-birth law).

The next section provides a description of the presence of English in Swedish society and the educational system, as well as an overview of some of the internally and externally motivated effects of trilingualism on low-prestige minority languages.

2.3 Trilingual Heritage Speakers

Because the bulk of research in the field of HL studies has centered on U.S. populations, the factor of multilingualism in HL use and maintenance has been largely irrelevant. HL research in multilingual settings should be concerned, among other things, with analyzing the cross-linguistic influence of societal languages on the HL, as well as how multilingualism affects HL
maintenance and use more generally. In the context of the Chilean community in Sweden, this study seeks to address these questions, focusing specifically on the impact of a *societally prestigious* L3 (English) on an immigrant minority L1 (Spanish).

There is a growing body of literature on the growth of English in European society (e.g., Cogo & Jenkins 2010; Pennycook 2014; Smit 2010). The remarkable reach of English includes domains as diverse as popular media, academia, business, and government. English is increasingly relied on as a *lingua franca* both within Europe and internationally, and has thus been implemented as a required subject of study often beginning at first grade. In Sweden, although young children experience exposure to English through television and music, they typically begin a formal study of it by around age eight (Cenoz & Jessner 2000). The status of English is especially advanced in northern Europe, where, according to Viereck (1996), it is ceasing to be viewed as a foreign language. Such a classificatory shift is linked to its increasing nativization, i.e., the emergence of a European English variety\(^\text{10}\) (Cenoz & Jessner 2000, 249; De Houwer & Wilton 2011; Forche 2012), and its role in the identity constructions of young Europeans (Gnutzmann, Jakisch & Rabe 2014). Recent research has also documented the spread of Content and Language Integrated Learning (CLIL) schools in Sweden (Yoxsimer Paulsrud 2014), which employ English-medium instruction.

In Sweden, English proficiencies among young adults as range along a continuum due to differences in experience abroad, opportunities for use, language aptitude, educational attainment, type of employment, and media exposure. Sharp (2001) used spoken corpus data to examine the use of English among young Swedes, and found their English to be generally “very good”. Despite their strong proficiency, Sharp explains that the language contact situation with English for young Swedes is mostly one-way or receptive, meaning that their chief contact with

\(^{10}\) Separate from U.K. varieties, which are also European.
English is via music, films, television and radio programs, computers and literature. Because of the indirectness of contact, meaning few opportunities to produce English output, and limited face-to-face interaction with native speakers, she concludes that their receptive abilities substantially exceed their productive skills. Accordingly, the second-generation Spanish speakers of this study all report at least some English proficiency (data to be presented in Chapter 4). The following is the typical language learning sequence for their demographic: L1 Spanish, early L2 Swedish, late L3 English. The terms *L1*, *early*, and *late* require operationalized definitions.

DeHouwer (2009) very strictly distinguishes between these terms. She defines an L1 as a language a child is exposed to from birth – not a day or a week after birth, but from the moment of birth\(^\text{11}\). She defines *Early Second Language Acquisition* (ESLA) as informal additional language learning, often in a daycare or preschool environment, when a child is between 1;5 and 4 years of age. *Late SLA* begins in a formal setting after age 6. She does not offer a definition for language learning that commences between birth and age 1;5, nor between ages 4 and 6. For children who are exposed to two L1s simultaneously, DeHouwer uses Meisel’s (1989) term *Bilingual First Language Acquisition* (BFLA), defined as “the development of two languages from birth in young children” (p. xiii). In consideration of these definitions, the general learning trajectory of the trilingual speakers of this study includes an initial stage of monolingual first language acquisition (MFLA), followed by ESLA of Swedish\(^\text{12}\), followed by the late acquisition of a third language (English).

\(^{11}\) There is recent evidence suggesting that some aspects of newborns’ cries (intonation) are acquired in utero (Mampe et al. 2009).

\(^{12}\) This is based on factors such as the high rate of Swedish proficiency among adult immigrants, Swedish language television programming, and the input most young children receive through the widely utilized government subsidized daycare/preschool. The participants' perspectives on their language trajectories were also taken into account.
As a relatively young area of study, most of the work in trilingualism research to date has dealt with third language acquisition (TLA) and the roles of the L1 and L2 in this process (e.g., Bocksay Pap 2015). Generally speaking, TLA has been found to be far more complex than SLA because of the dual filter through which speakers process new linguistic information. Several studies have found the role of the L2 in TLA to be more important than originally thought (Carvalho & Bacelar da Silva 2006; Cenoz, Hufeisen, & Jessner 2003; Rothman 2010).

The reverse effect – that is, of the L3 on L1 or L2 development – has received much less attention. Goorhuis-Brouwer and de Bot (2010) addressed the common preoccupation in the Netherlands that early English learning (defined as learning taking place in the early school years) negatively impacts L1 Dutch development. Their findings showed no measurable lag in Dutch acquisition for Dutch-background L2 or non-Dutch background L3 English learners. However, they note the little that is known on how L3 English learning affects the minority L1 development of young English language learners in Europe (p. 300). This is particularly true in the case of low-prestige immigrant minority languages, as research has pointed only to the non-detrimental effects of English learning on relatively high-prestige regional minority languages. Cenoz (2005), for instance, found that the early introduction of L3 English did not negatively affect L1 Basque or L2 Spanish development when compared with a control group of non-English learners. Even positive effects on the L1 are observed when more specific aspects of proficiency (such as metaphorical density) are tested (Kecskés & Papp 2000). On the unfavorable position of HLs in spite of ideologies of multilingualism in Europe, de Bot and Gorter (2005, p. 614-615) share the following:

At the European level, policies aimed at language maintenance have been limited to RM [regional minority] languages…The value of knowing more languages is generally accepted, but this acceptance does not usually extend to RM or IM [immigrant minority] languages. The focus is mainly on English, which has grown in importance over the last decades and which, in many
(smaller language) countries in Europe, is now more and more becoming the language of higher education, trade, and commerce. Attitudes towards English are generally positive, and the need to know English is pervasive in the school-going population in European countries (Bers & de Bot, 2005). So, there is a paradox in the sense that multilingualism is generally seen as an asset when it concerns English as a second or third language but not when it concerns minority languages.

Clyne (1997) provides one perspective of L3 influence on the lexicon of immigrant minority L1s. In an Australian-based study of bilinguals-turned-trilinguals speaking various combinations of languages (but with English as the common L3), Clyne notes that there was transfer of English lexical items into the other languages, and that these items functioned as a bridge not only between English and the speakers’ L1 or L2, but between the L1 and L2 as well. Although this study dealt with a scenario in which the L3 (English) was the majority language, the implications of Clyne’s findings for the Swedish context seem to include the following: (1) English as a high-status L3 can act as a direct source of lexical transfer to the participants’ L1 Spanish, and (2) the speakers’ L3 English can also indirectly impact their L1 Spanish by way of their L2 Swedish. The latter effect may be particularly strong given that the structural and lexical similarities between English and Swedish create a condition found to compound cross-linguistic influence (Jarvis & Pavlenko 2008; Ringbom 2007).

On the one hand, Clyne (1997) and Hoffman (2001) have observed that when speakers acquire a prestigious L3, they tend to begin using it frequently, partially abandoning one of the remaining two languages (typically the one of lowest societal status). This suggests that the early acquisition of L3 English by SHS in Sweden may contribute to a more rapid loss of their HL, given its inferior status to Swedish. On the other hand, additional work has shown that multilingualism can actually aid L1 maintenance through its benefits of enhanced metalinguistic awareness and metaphorical/conceptual competence (Cenoz 2000; Kecskés & Papp 2000). Thus, it is possible that these benefits are reflected in the lexical knowledge of HS. The present study
tests the relationship between the G2 participants' level of English proficiency and HL maintenance (lexical knowledge).

2.4 Spanish Heritage Speaker Lexicon

2.4.1 Defining vocabulary and lexical knowledge

According to Alderson (2005, p. 88), “language ability is to quite a large extent a function of vocabulary [knowledge]”. The construct of vocabulary knowledge is complex, however, involving several dimensions. Hence, the difficulty of the question “What do you know when you know a word?” In short, “lexical knowledge is made up of different kinds of word knowledge, and not all can be mastered simultaneously” (Schmitt 2010, p. 22). Schmitt explains that “not only is word learning incremental in general, but learning of the individual word knowledge aspects as well” (p. 23).

Anderson and Freebody (1981) introduced the basic distinction of breadth and depth of word knowledge; breadth of word knowledge refers to the number of words a speaker-learner knows, and depth of word knowledge is what is known about these words. Meara (1996) proposed a third dimension involving the ease with which speakers can access and produce words. Daller, Milton, and Treffers-Daller (2007) label this third dimension fluency of knowledge. According to Milton and Fitzpatrick (2014), this three-dimensional framework facilitates the distinction between learners who know many words and a lot about these words but encounter difficulty using them, and learners “who can quickly and naturally activate this knowledge for communication” (p. 7).

Lay understandings and even language teaching generally focus on the basic form-meaning relationship of word knowledge (how breadth tends to be measured), yet other dimensions are required for a complete mastery, or a higher “quality of understanding”
(Anderson & Freebody 1981, p. 92-93) of a lexical item. On what the concept of depth of word knowledge entails, Read (2000, p. 211) outlines the following three (interrelated) components:

1. **Precision of meaning**: The difference between having a limited, vague idea of what a word means and having a much more elaborated and specific knowledge of its meaning.

2. **Comprehensive word knowledge**: Knowledge of a word, not only its semantic features but also orthographic, phonological, morphological, syntactic, collocational, and pragmatic characteristics.

3. **Network knowledge**: The incorporation of the word into a lexical network in the mental lexicon, together with the ability to link it to — and distinguish it from — related words.

Because the various aspects of depth of lexical knowledge are difficult to isolate for measurement purposes, those in research and teaching contexts often opt for the simpler distinction of receptive versus productive knowledge (related to the third dimension in Meara’s 1996 framework). According to Melka (1997), receptive and productive mastery fall on a continuum, with knowledge progressing from receptive to productive as learning advances. Alternatively, Meara (1997) suggests that the shift from receptive to productive mastery of a lexical item corresponds to its level of integration into the mental lexicon, i.e., the number of links it has with other productive items. Such an organizational perspective in which linked items play a large role in word acquisition and activation “has the potential to explain how students can learn some words productively with very little input over a short period of time,” or why the same words "sometimes seem to be known productively and at other times do not” (Schmitt 2010, p. 81). According to these theories, productive mastery requires a deeper knowledge of a given lexical item than does receptive mastery; hence its utility as a proxy for more detailed assessments of depth of lexical knowledge.
The question “What is a ‘word’?” has elicited much scholarly debate. A word is attributed different characteristics depending on the theoretical approach and the descriptive context. Though numerous, definitions across disciplines generally fall within three main criteria: acoustic and semantic identity, morphological ability, and syntactic mobility. The Routledge Dictionary of Language and Linguistics (2006) provides the following examples of definitions for the concept of “word”:

(a) phonetic-phonological level: words are the smallest segments of sound that can be theoretically isolated by word accent and boundary markers like pauses, clicks, and the like

(b) orthographic-graphemic level: by blank spaces in writing or print;

(c) morphological level: words are characterized as the basic elements of grammatical paradigms like inflection and are distinguished from the morphologically characterized word forms, cf. *write* vs *writes, wrote, written*; they are structurally stable and cannot be divided, and can be described as well by specific rules of word formation;

(d) lexical-semantic level: words are the smallest, relatively independent carriers of meaning that are codified in the lexicon, and

(e) can be described syntactically as the smallest permutable and substitutable units of a sentence.

Because of these distinct conceptualizations, the term *word* is often discarded by linguists in favor of concepts like *morpheme, lexeme, lemma, lexical unit*, and $X^o$ (in X-bar theory).

At a very basic level, a “lemma”, is a base (*trabajar*) from which its inflected forms derive (*trabajo, trabaja*, etc.). Related to this is the notion of a “word family,” a set of words that share a common root meaning (*trabajar, trabajo, trabajador*, etc.). In studying vocabulary knowledge, not all words are equal; function or grammatical words (e.g., prepositions, articles) versus content or lexical words (e.g., nouns, adverbs) are different and reflect different stages of acquisition. Accordingly, in text analyses (including the present study) where vocabulary is the
linguistic variable of study, a distinction is made between tokens (i.e., the total number of words) and types (i.e., the total number of different words).

Because of the problems associated with the fuzzy notion of a “word”, the more inclusive terms lexeme, lexical unit or lexical item are often used in vocabulary research. These three interchangeable terms are all understood as “an item that functions as a single meaning unit, regardless of the number of words it contains” (Read 2000, p. 1). Thus, as Read (2000) shows, the following six examples are lexemes with a shared meaning: die, expire, pass away, bite the dust, kick the bucket, give up the ghost. Accordingly, Daudaravicius (2010) explains that lexical units can be (1) a single word; (2) the habitual co–occurrence of two words; and (3) a frequent recurrent uninterrupted string of words (i.e., a collocational chain). In this study, the terms lexical item and lexical unit are favored in the description and analysis of the participants’ productions (consisting of both single and two-word units) in a lexical availability task.

Vocabulary has been the subject of increasing focus in the classroom. In second language education, there has been a fairly recent transition from a grammar-based approach to teaching, in which grammar rules are learned and subsequently applied to examples, to a focus on the learning of high-frequency multiple word “chunks” of language that aid learners in gradually identifying larger patterns (Nattinger & DeCarrico 1992; Ellis 2002). Eyckmans (2004) discusses the increasing lexical focus in language teaching starting in the 1990s, citing the novel emphasis on the notion of “grammaticalised lexis” in place of “lexicalized grammar” in Lewis’s (1993, 1997) teacher training manuals. Similarly, Read (2000) explains how “after a lengthy period of being preoccupied with the development of grammatical competence, language teachers and applied linguistic researchers now generally recognise the importance of vocabulary learning and are exploring ways of promoting it more effectively” (p. 2).
In fact, Anderson and Freebody (1981, p. 77) contend that “measures of vocabulary knowledge are potent predictors of a variety of indices of linguistic ability.” As it is known that HL acquisition patterns after both L1 and L2 learning in different ways, it stands to reason that the import of vocabulary learning for global language development holds in HS population as well. According to Ellis (1997), lexical proficiency is critical for the acquisition of grammar because comprehension of the discourse facilitates the transparency of grammatical structure and grammatical relationships. Nation (1990, p. 20; 2001) has called for “a systematic, principled approach to vocabulary development”, which he argues leads to more successful language learning. In the realm of HL teaching, obtaining an empirically based understanding of the nature of HS lexicon in a given population is critical to implementing such a “systematic, principled approach.”

Studies in both L1 and L2 populations, mostly focusing on English, have found correlations between vocabulary and specific language skills. One example is the correlation between lexical knowledge and reading skills (Laufer 1989, 1992; Schmitt, Jiang, and Grabe 2008; Wittrock, Marks, & Doctorow 1975). Studies of HS have reported similar correlations, particularly in research on receptive lexical knowledge. A series of studies by Polinsky (1997, 2000, 2006), for example, found a strong correlation between the receptive lexical knowledge of Russian HS in the U.S., measured by a 200-item recognition task, and oral abilities in the areas of agreement, case marking, aspectual and temporal marking, pro-drop, co-reference, and embedding. Fairclough (2011) has also found correlations in this regard for SHS using a similar lexical recognition task and multiple proficiency measures. Benmamoun et al. (2010) discuss the promise of this emerging relationship: “If structural attrition and lexical proficiency are correlated, lexical proficiency scores, which are relatively easy to obtain, can serve as a basis for
the characterization and ranking of incomplete learners on a continuum (rather than in a discrete group)” (p. 22). Among educators and researchers working with HS populations, the option of quick yet reliable lexical assessments could come as a welcome alternative to lengthy grammar tests and oral interviews.

2.4.2 The SHS lexical repertoire

I now turn to research on the specific characteristics of SHS lexicon. As discussed in the introduction, despite its significance in language learning as well as its many complexities, little is known about the long-term lexical development of early bilinguals, including HS (Genesee 2006). In the context of SHL, Montrul (2012, p. 13) underscores the “highly unexplored” status of lexical knowledge. To date, the bulk of research on the linguistic systems of SHS has focused on morphosyntactic features, such as tense and aspect (Silva-Corvalán 1994; Montrul 2002), mood (Lynch 1999; Montrul 2007; Silva-Corvalán 1994), pronoun usage (Silva-Corvalán 1994; Otheguy & Zentella 2012), object marking (Montrul & Bowles 2009), and gender (Montrul et al., 2008), as well as discourse markers (Torres & Potowski 2008a).

One recurrent finding throughout the limited body of research on SHS lexicon is that, due in part to lack of formal instruction in the HL, these speakers exhibit difficulties with the lexical aspects of academic or formal language during production, often using borrowed, colloquial, stigmatized, and high frequency words in these contexts (Achugar 2003; Sánchez Muñoz 2010; Valdés & Geoffrion-Vinci 1998). Nevertheless, these studies also show that SHS do in fact make some lexical and other stylistic adjustments to their language in response to variations in situational formality.

It is important to note that retrieval problems rather than lack of lexicon are believed to account for some of the lexical difficulties HS face. Polinsky and Kagan (2007), for example,
reported slower speech rates for Russian HS, which they attributed to difficulties with lexical access. In addition to diminished speed and accuracy in lexical access (Schmid & Köpke 2009), complete lexical loss can occur (Schmid 2007). But determining whether lexical or other HL phenomena are a result of incomplete acquisition or attrition is notoriously difficult (Benmamoun et al. 2011). In the next section I will review the topic of lexical access in the context of age of acquisition (AoA) effects on HS lexical knowledge. In addition to studies focusing on AoA, other HS vocabulary research has examined receptive knowledge, production, and lexical transfer. I will now review some key studies within these areas, with a predominant focus on SHS.

2.4.3 Research in SHS lexicon

AoA

The effect of AoA on lexical access is one area that has garnered the attention of those studying attrition (e.g., de Bot 1998; Schmid & Köpke 2009). One study that lends insight into the receptive and productive aspects of AoA is that of Hulsen (2000), who administered picture-naming and picture-word matching tasks in Dutch and English to three generations of Dutch speakers in New Zealand. For all speakers, the picture-naming (productive) task proved more difficult than the picture-word matching (receptive) task. However, performance on the picture-naming task decreased more markedly across generations, i.e., speed and accuracy in picture naming were significantly different between the first and second/third generation participants. This showed that production is affected more than comprehension in HL attrition, and that speakers acquiring a majority-group L2 at an early age are more vulnerable to L1 lexical loss than adult immigrants.
Montrul & Foote (2014) used both the AoA of individual lexical items and of the Spanish language itself to compare the speed and accuracy of lexical access in "proficiency-matched" HS and L2 learners of Spanish. They classified lexical items in the following manner: Early L1-Early L2 (lápiz 'pencil', besar 'to kiss', cansado 'tired'); Early L1-Late L2 (cuna 'crib', barrer 'to sweep', travieso 'mischievous'); and Late L1-Early L2 (correo 'mail', resumir 'summarize', soltero 'single'). Both participant groups performed a lexical decision task in Spanish and an English-Spanish translation decision task. With respect to the global AoA of the L1 and L2, the results showed no accuracy differences in lexical access between the two groups, but speed did play a significant role in that the L2 learners responded more quickly than the HS. With respect to word AoA, differences were noted between the groups: HS displayed more speed and accuracy than L2 learners with words learned early in L1 Spanish, while the L2 learners were faster and more accurate than HS with words learned early in L2 Spanish. The authors also considered the role of lexical class in participant results, finding that both groups favored nouns over verbs and adjectives in terms of speed and accuracy (though the effect did not reach statistical significance). The results thus suggest that maturational constraints (i.e., global AoA) do not affect visual word recognition. Also, the effect of word AoA (i.e., the greater retention of early-acquired words by both groups) showed the important role experience and use plays in lexical access.

**Receptive vocabulary**

Because receptive abilities generally exceed productive abilities, research aiming to establish the outer limits of SHS vocabulary has made use of receptive tasks. Fairclough (2011, 2013a) explored SHS receptive vocabulary knowledge through tasks based on corpus frequency data. In her first study, she used a 10-minute computerized lexical recognition test designed for an SHS
university course placement test. The test was elaborated using Davies’ (2006) Spanish frequency dictionary and the online *Corpus de Referencia del Español Actual* (CREA) database. A cross-sectional group of 330 students performed a lexical decision task on 120 words representing various levels of frequency. A total of 24 words were selected randomly within each of the five 1000-word frequency bands of the 5,000 most frequent words listed in Davies’ (2006) Spanish frequency dictionary. To control for guessing, 80 pseudowords were added to the 120 randomly selected words. The students were presented with a single word stimulus to which they were to mark “yes” if they could explain its meaning to a friend.

Significant correlations were found between the results from the lexical task and both a cloze test and "multiple-task" test, pointing to the reliability of the measure in distinguishing proficiency levels. Crucially, however, the correlation only held significance for the lower proficiency students. For speakers of advanced proficiency, Fairclough (2011) concluded that the range of words would need to be expanded “to avoid the ceiling effect” (p. 17), i.e., a similar result for all advanced proficiency students. The high performance of the advanced speakers was surprising given that the 5,000 word frequency limit was chosen in consideration of the following reports by Davies (2005, p. 110): (1) a frequency list of about 4000 total words covers about 90% of all words heard in a typical [monolingual] conversation, and (2) about 7000 words are needed to achieve 90% coverage in fiction writing; and nearly 8000 for non-fiction writing. In other words, the level of the 5000 most frequent words in Davies’ corpus “appeared to be situated on the limit between words of high and low frequency in Spanish” (p. 291). But in Fairclough’s results, many of the HS students exceeded expectations with regard to receptive knowledge of Spanish vocabulary, such that a test limited to words within the first 5 frequency bands (i.e., 0 – 5,000 words) would not effectively distinguish linguistic ability among high
proficiency HS. A subsequent report (Fairclough 2013a) showed that the receptive SHS students who were included in the research recognized up to 60% of the words representing the first 5 frequency bands, while the participants of higher proficiency recognized around 90% of the words (hence, the ceiling effect reported in Fairclough 2011). The limited proficiency SHS students were found to have a receptive lexicon of up to 3,000 words.

**Lexical production**

The limited research on SHS productive lexicon has also found it to be quite strong, despite its greater vulnerability to attrition and access/recall weakening. In section 2.4.2 I cited studies that have looked at the stylistic range of SHS and discussed what those findings say about the lexicon. That work, although based on recorded spoken data collected in the variable contexts of class presentations, interviews, and casual conversations, was topically limited, and did not consider how additional language-external variables influence lexical knowledge and use in the HL.

Other work on the lexical production of SHS has focused on lexical availability. Given that lexical availability is central to the theoretical-methodological framework of my study, these studies are particularly relevant to its design and data analysis. An overview of the research tradition of lexical availability, including its special significance in Hispanic dialectology, is provided in Section 2.5. Structured around a variety of semantic domains, lexical availability tasks measure the number and types of written words speakers produce in each thematic area within a limited period of time. The lexical availability of native Spanish speakers in Spain and Latin America has been widely researched, and has included several studies carried out in language contact situations (reviewed in section 2.5). In the context of the bilingual
Hispanophone population in the U.S., three studies have used this methodology: Cooper (1971); Victery (1971); and Moreno-Fernández (2007).

Working with bilingual Puerto Ricans in New Jersey (generation unspecified), Cooper (1971) used a word association (lexical availability) task to study the participants’ use of English and Spanish in the following social domains: family, neighborhood, religion, education, and work. Participants were asked to produce within one minute as many words as they could that were related to the specific topic. Cooper found that the participants produced a greater number of items in the language that they normally used in each domain, showing the important role of experience in lexical fluency.

Victery (1971) studied the Spanish or English vocabulary elicited in ten different subject areas by 100 adolescents representing the following groups (1) Spanish monolinguals in Monterrey, Mexico, (2) English monolinguals from Houston, Texas, and (3) bilinguals from Houston, Texas. Victery did not control for education or SES. Interestingly, in terms of total number of types (i.e., different words), the bilingual group produced more than the monolingual groups. With respect to total lexical output, the bilingual participants produced a higher average of lexical items than the Spanish monolinguals, but a lower average than the English monolinguals, in all 10 subject areas. With regard to the specific semantic domains covered, the bilingual participants’ productions were most numerous in these categories (in descending order): animales 'animals', la ciudad 'the city', guerra 'war', familia 'family', música 'music'. Fewer items were produced under the categories of Dios 'God', medicina 'medicine', trabajos y profesiones 'work and professions', espacio 'space', diversión 'entertainment'. Victery also found that the female productions in each group outnumbered the male productions. His results reflect the robust Spanish lexicon of bilingual adolescents in an English dominant society.
Moreno-Fernández (2007) conducted a study of the lexical availability of Spanish-English bilingual adolescent Hispanics in Chicago. His participants, students enrolled in two course levels of high school Spanish for HS, provided word lists in response to 20 semantic prompts (listed in Table 2-3). The first sixteen of these prompts coincide with the “centers of interest” used in the Proyecto Panhispánico de Léxico Disponible, whose objective was to establish the lexical norms of the diverse areas of the Spanish-speaking world (López Morales 1999). The remaining four consisted of the following fields: los colores 'colors'; vida y familia 'life and family'; salud y enfermedades 'health and sicknesses'; árboles y plantas 'trees and plants'.

<table>
<thead>
<tr>
<th></th>
<th>Semantic Prompts Used in Moreno-Fernández (2007)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cuerpo humano</td>
<td>‘The human body’</td>
</tr>
<tr>
<td>2</td>
<td>La ropa</td>
<td>‘Clothing’</td>
</tr>
<tr>
<td>3</td>
<td>Las partes de la casa</td>
<td>‘Parts of the home’</td>
</tr>
<tr>
<td>4</td>
<td>Los muebles de la casa</td>
<td>‘Furniture in the home’</td>
</tr>
<tr>
<td>5</td>
<td>Comidas y bebidas</td>
<td>‘Food and drink’</td>
</tr>
<tr>
<td>6</td>
<td>Objetos colocados en la mesa para la comida</td>
<td>‘Objects placed on the table for a meal’</td>
</tr>
<tr>
<td>7</td>
<td>La cocina y sus utensilios</td>
<td>‘The kitchen and utensils’</td>
</tr>
<tr>
<td>8</td>
<td>La escuela (muebles y materiales)</td>
<td>‘School (furniture and materials)’</td>
</tr>
<tr>
<td>9</td>
<td>Calefacción, iluminación y medios de airear un recinto</td>
<td>‘Heating, lighting, and methods of ventilating a space’</td>
</tr>
<tr>
<td>10</td>
<td>La ciudad</td>
<td>‘The city’</td>
</tr>
<tr>
<td>11</td>
<td>El campo</td>
<td>‘The countryside’</td>
</tr>
<tr>
<td>12</td>
<td>Medios de transporte</td>
<td>‘Means of transportation’</td>
</tr>
<tr>
<td>13</td>
<td>Trabajos del campo y del jardín</td>
<td>‘Outdoor jobs and yardwork’</td>
</tr>
<tr>
<td>14</td>
<td>Los animales</td>
<td>‘Animals’</td>
</tr>
<tr>
<td>15</td>
<td>Juegos y distracciones</td>
<td>‘Games and entertainment’</td>
</tr>
<tr>
<td>16</td>
<td>Profesiones y oficios</td>
<td>‘Professions and vocations/trades’</td>
</tr>
<tr>
<td>17</td>
<td>Los colores</td>
<td>‘Colors’</td>
</tr>
<tr>
<td>18</td>
<td>Vida y familia</td>
<td>‘Life and family’</td>
</tr>
<tr>
<td>19</td>
<td>Salud y enfermedades</td>
<td>‘Health and diseases’</td>
</tr>
<tr>
<td>20</td>
<td>Árboles y plantas</td>
<td>‘Trees and plants’</td>
</tr>
</tbody>
</table>
Moreno-Fernández’s analysis differed from the goals of the Proyecto Panhispánico, however, in that it focused on the presence of Anglicisms in the students’ productions. Of the participant groups’ 20 most available words under each category (400 total items), he found that only 6.5% could be considered English-influenced lexicon. This percentage rose to 15% when considering the overall lexical output, which included a large number of idiosyncratic items with low availability indexes (i.e., items produced relatively late in the lists by few participants).

Importantly, Moreno-Fernández (2007) found that a greater number of Anglicisms were produced under topics that had little to do with the participants’ day-to-day experiences (e.g., juegos y distracciones 'games and entertainment'; calefacción, iluminación y medios de airear un recinto 'heating, lighting, and methods of ventilating a space’). Additionally, of the independent variables considered — gender, generation, and level of Spanish (measured by course level) — the latter presented the strongest correlation with the production of Anglicisms. In other words, those who were deemed the least proficient in Spanish offered the most English-influenced lexicon.

Although his study was concerned with contact lexicon, Moreno-Fernández (2007) also provided the total number of word types offered by the entire participant group in each category. Although the participant exhibited robust knowledge in all 20 domains, the top five themes for which most vocabulary was produced were: comida y bebida 'food and drink; vida y familia 'life and family'; la ciudad 'the city'; animales 'animals'; juegos y distracciones 'games and entertainment'. These results coincide significantly with those reported in Victery (1971), save in the case of juegos y distracciones/diversión, a domain in which Victery’s participants produced relatively few items. Like Fairclough (2011), Moreno-Fernández notes the extensive vocabulary knowledge SHS display in their HL. In addition to a limited reliance on English language
borrowings, Moreno-Fernández’s results show the speakers’ productive (i.e., advanced) abilities across diverse, albeit rather concrete, thematic areas.

Given that Moreno Fernández (2007) analysis was narrowly focused, important questions remain. His study’s quantification of Anglicisms (rather than a general type-token analysis of the body of words) constitutes an indirect, and perhaps somewhat problematic, measure of lexical knowledge. Due to his focus on Anglicisms, the relationship between the independent variables (gender, generation, and level of Spanish) and the participants’ productive lexicon was not included in the results. Another issue has to do with the relationship between global proficiency and lexical knowledge, and the need for an external proficiency measure. In his study, Moreno-Fernández measured global proficiency according to class level. This is questionable because the methods of placement are uncertain and uncontrolled, and thus difficult to justify.

A further issue has to do with the lack of correlation Moreno-Fernández (2007) found between the variable of generation and English-influenced lexicon. This result may be skewed due to the criterion used for inclusion in the first generation (i.e., place of birth), which does not mirror the commonly employed cut-off of post-pubescent arrival in the host country (Silva-Corvalán 1994). Because many of those born abroad had presumably arrived at an early age (AoA was not provided), the division and categorization of generations in this manner (i.e., born in Latin America vs. born in the U.S.) is fraught. Under this classification, “Born in Latin America” could mean that one arrived in Chicago at 6 months or that one arrived at age 13, implying very different language trajectories and amount of exposure to English. Given that cross-generational analyses on the acquisition and change of migrant languages have been shown to benefit from knowing (1) age of arrival and (2) number of years in the host country, it would
be likewise important to consider these in lexical studies with generation as a variable. Relatedly, Moreno-Fernández presented his results in isolation of previous work on the lexical availability of monolingual Spanish speaking groups. While bilingual and monolingual communities of speakers are known to differ in important ways, contextualizing bilingual lexical behavior within comparable monolingual findings can provide important insights, particularly with regard to the attrition experienced by the first generation of immigrants.

In addition to the need for research on SHS productive vocabulary to address the aforementioned areas of appropriate generational divisions and independent proficiency measures, work aimed at a comprehensive examination of the domain-based lexical knowledge of SHS would benefit from the inclusion of categories representing more abstract, complex, or societal topics, such as la política ‘politics’ or problemas sociales ‘social problems’. That is, the elicitation of vocabulary across a broad range of domains from generationally distinct speaker groups is needed to gain a fuller picture of intergenerational (dis)continuities in the quantitative and qualitative aspects of vocabulary knowledge. This type of data can shed greater light on the lexical repertoires of HS and their ability to participate in various real-world contexts. In relation to the first generation, a domain-based analysis can lend insight into the areas most vulnerable to attrition or change.

**Lexical transfer**

There is a large body of work dedicated to the study of lexical transfer in the context of English-Spanish bilingualism in the U.S. (Lipski 2008; Mendieta 1999; Otheguy & García 1993; Otheguy, García, & Fernández 1989; Poplack, Sankoff, & Miller 1988; Sánchez Munoz 2010; Smead 1998; Zentella 1997; Varra 2013; among many others). Despite the diverse backgrounds and proficiencies of HS, transfer from the majority language is usually present to some degree in
their oral and written production. Fairclough (forthcoming) explains how “mixing or alternating
the two languages are options that bilinguals have, provided that certain functional, structural,
semantic and communicational constraints are observed” (p. 10). Lexical transfer is also attested
in the Spanish of adult migrants, including Otheguy & García (1993); Poplack, Sankoff & Miller

According to Fairclough (forthcoming), lexical transfer typically appears in the form of
code-switching and borrowing (including loanwords, phrasal calques, and semantic extensions),
and lexical creations (Sánchez Muñoz 2010). Code-switching, the alternation between two
languages in continuous speech, can be in the form of single lexical items or entire phrases that
maintain the original phonology. Loanwords are understood as items whose form and meaning
are borrowed into a language, generally undergoing phonetic and/or morphological adaptation
(marqueta ‘market’). Calques are phrases that contain receptor language words but copy donor
language structure (correr para presidente ‘to run for president’). Semantic extensions are
similar to calques in that existing forms incorporate new/additional meanings due to contact with
cognates or partial translation equivalents in the majority language (e.g., using aplicar in the
sense of applying for employment). Finally, lexical creations, like loans, involve the transfer of
form and meaning, but differ in that they are infrequent, “often belonging to an individual’s
idiolect (e.g., endurar ⇒ ‘to endure’ = aguantar)” (Fairclough, forthcoming, p. 11). These types
of borrowings are referred to elsewhere as nonce or spontaneous borrowings. Otheguy and
Garcia (1993) employ the term “contact neologisms” to refer to all of the above types of lexical
borrowing.

With specific regard to lexical transfer in SHL, I have already reviewed Moreno-
Fernandez’s findings on the presence of English-origin borrowings in a written lexical
availability task. Fairclough (2006) provides an analysis of the different types of transfer phenomena in SHS university placement exams, a context in which the sole use of Spanish was expected. The data consisted of 150 writing samples completed by SHS in order to place into one of five different levels in the Spanish program at the University of Houston. The study examined the types and extent of English-Spanish interaction in four brief essays, whose topics elicited different kinds of discourse: descriptive, narrative, argumentative, and hypothetical. Study results revealed low levels of any kind of English transfer (7.15 words per average sample of 308 words). Similar to the outcome in Moreno-Fernández (2007), the influence of English was stronger in the writings of less proficient students. Phrasal calques were the most frequent type of transfer across all levels, but Fairclough (2006) notes their insignificance when compared to the total number of words produced (total number of words: 46,244; total number of phrasal calques: 751). The other phenomena quantified, switches to English, loans, and single-word calques (i.e., semantic extensions), were negligible.

Similarly, Garza (2013) analyzed the oral and written academic language of SHS students divided across four groups according to proficiency (determined by course level). The author found that there was an inverse relationship between lexical transfer and proficiency; i.e., the higher the level of Spanish, the fewer instances of lexical transfer. The relationship between higher levels of English proficiency and borrowing frequency has been documented elsewhere (Otheguy et al. 1989; Torres & Potowski 2008b; Varra 2013). However, this work has also confirmed that the presence of loanwords is by no means limited to SHS repertoires. Given that borrowing serves multiple purposes, including filling culturally motivated lexical gaps, and preventing misunderstandings or increasing neutrality among Latinos of different national origins (Potowski 2011; Zentella 1990), it is a mechanism first generation speakers rely on as
well. More research is needed to understand how, beyond simple quantitative distinctions, borrowing behavior differs among speakers of distinct profiles. Examining borrowing across a broad array of semantic domains and the quantity as well as quality of borrowings across generations can lend additional understanding of this aspect of lexical change.

Semantic domains were indirectly involved in a study carried out by Sánchez Muñoz (2010), which examined lexical transfer in the oral production of SHS across the contexts of: (1) class presentations, (2) interviews, and (3) conversations. Among other lexical features, she considered (1) Single word switches or English fillers (e.g., *la* deadline; you know), and (2) lexical creations (nonce borrowings) in Spanish based on English words (e.g. *endurar* 'to endure'). In terms of switches to English, these were most prevalent in conversations and least prevalent in presentations. Alternatively, the number of lexical creations was highest in the presentations and lowest in conversations. This suggests that in less formal settings, heritage speakers maximize communicative efficiency by employing borrowings of common usage and switching between Spanish and English, while in more formal settings they attempt to transfer their knowledge of low frequency, formal register words in English to Spanish. The lexical variation across registers displayed by these participants refutes the notion of a monostylistic variety spoken by SHS, showing, rather, that they make linguistic adaptations using the resources available to them.

Fairclough (2013b) studied whether certain single and multiple-word high-frequency calques were in the process of lexicalization in U.S Spanish. These included, for example, “solicitud / aplicación” (*application*), “inscribirse / registrarse” (*register*), “universidad / colegio” (*college*), and “tomar una decisión / hacer una decisión” (*make a decision*). Fairclough based her analysis on 101 three-paragraph English-Spanish translations completed by Hispanic

---

13 The author did not indicate how she determined whether a borrowing was idiosyncratic (nonce).
university students of various proficiency levels. Although the students had previously been taught corresponding forms of wider usage, the results indicated that a significant proportion of the participants employed the calques in question, supporting the idea that these forms have undergone lexicalization.

The varied results of studies on SHS lexicon—from the rich receptive lexicon of advanced HS in Fairclough (2011), to the register variation attested in Sánchez Muñoz (2010); from the minimal English influence in Moreno-Fernández (2007) to the primarily culturally-based borrowing reported in Gamboa González (2003)—suggest that the claim that adult SHS exhibit a lexical restriction to “home” and “family” domains is overly simplistic. But further clarification is needed with regard to the notion of restricted domains. What domains ought to be considered basic? And, in a given speaker population, to what degree is the lexicon of SHS skewed toward basic domains? Are there notable areas in which SHS lexicon is especially sophisticated, and how does this relate to community, family, or individual background? For example, as mentioned in the introduction, there is evidence to suggest that the exile background of the Chilean diaspora has directly impacted the ability of the second generation to converse about topics in politics (Becker 2013; Gibbons & Ramirez 2004). At the same time, SHS of other heritages may exhibit lexical strengths in different areas.

While the research reviewed in this section points to the need for a fuller understanding of SHS lexical abilities, to complement work on the morphosyntactic aspects of HS grammar, the distinctively context-dependent, broad, and open nature of the lexicon presents methodological challenges. The methodology employed in lexical availability research represents a useful approach to synchronically evaluating the words available to SHS for communication on a
variety of topics. In the next section I provide an overview of the research tradition of lexical availability, whose productivity in Hispanic linguistics is significant.

2.5 Lexical Availability

Available lexicon is understood as the sum of lexical units speakers have in their mental systems that can be recalled and produced in the treatment of a particular topic (Moreno-Fernandez, 2007). Tests of lexical availability measure the ease with which speakers can produce words as members of a given semantic domain (e.g., ‘the city’; ‘clothing’, ‘politics’), the types of words produced, and the prevalence that some words have over others in production. This is accomplished through the administration of timed word association tasks that elicit vocabulary belonging to a variety of semantic domains. Such data provide a representation of the available lexicon of a particular Spanish-speaking zone or group of speakers (e.g., youth, elderly, socioeconomic class, females, etc.). Inspired by methods in traditional neuroscientific research, word association tasks have aided language practitioners in a number of research and instructional objectives.

Prior to the elaboration of the concept of lexical availability in the early 1950s, lexicographers and other researchers using quantitative and statistical measures were primarily concerned with determining the “basic” lexicon of a given language, which they accomplished through general frequency counts using text-based corpora. The basic lexicon of a language variety was understood as the lexicon known by all speakers of that variety, and was intended to function as a resource in the elaboration of materials for second language learners.

However, researchers eventually became perplexed by the fact that many commonly known and used words did not appear in the frequency counts (Samper Padilla et al. 2003). The words absent from the lists were familiar to all speakers, but were characterized by an overall
low frequency in the language variety, which, according to the established criteria, did not qualify them for inclusion in its basic lexicon. An example of this problem is found in the *Diccionario Básico del Español de México* [Basic Dictionary of Mexican Spanish], where such familiar words as *cabello* [hair], *enfermera* [nurse] and *motocicleta* [motorcycle] are absent (Alba 1995a). This discovery served to show that there is no inherent correlation between the corpus frequency of a lexical item in a given language and its familiarity or utility for speakers. Because of the relative subjectivity of corpora, it became clear that frequency alone could not establish which words form part of the basic lexicon (i.e., words known and used by all speakers) of a language variety, but rather frequency contextualized within specific semantic domains. In other words, researchers aimed to find out the frequency of different words within the context of a particular topic.

In order to more effectively determine the fundamental lexicon of their language, French researchers developed a method of data collection that relied on the use of controlled association tasks capable of eliciting thematic lists of lexical items (Michéa 1953; Gougenheim, Michéa, Rivenc, & Sauvageot 1964). These pioneers of the study of lexical availability, motivated by the desire to facilitate the acquisition of French, developed a group of some 16 thematic stimuli meant to cover a significant portion of the basic human experience, including ‘Body Parts’, ‘Clothing’, ‘Professions’, ‘Food and Drink’, etc. Through the generation of lexicon within predefined semantic categories, or “centers of interest”, this method would account for both the frequency and familiarity of words. That is, word frequency would continue to be measured, but within the circumscription of a particular domain.

Again, as tested by this method, “available” words are those that immediately appear in the mind of the speaker when a given topic is presented. López Morales (1999) describes
available lexicon as that which is recalled and realized when certain associations are produced (p. 11). Related to the concept of lexical availability are Rosch’s (1973, 1975, 1978) prototype theories, which involve the notion that some members of a semantic category are more fitting or salient than others based on the prominence of their characteristics. To give a widely cited example from Rosch’s work, a robin is typically perceived as a more central member of the bird category than is, say, the less “birdy”, and thus more peripheral, penguin. In other words, often one or more prototypes can be identified for a given category (in addition to ‘Birds’, Rosch looked at ‘Vegetables’, ‘Fruit’, ‘Furniture’, ‘Clothing’, ‘Carpenter’s Tools’, among others), to which the characteristics of other category members are compared (below the level of consciousness) to determine their relative exemplarity of the category. In a lexical availability task, the more readily a group of speakers produce a lexical item in response to a category prompt, the higher its prototypical and communicative value appear to be for the community/group in question. In this way, data obtained through lexical availability tasks provide insight into the associations between prompts and the words produced, as well as among the words themselves (i.e., semantic networks). Crucially, the conceptualization of categories can vary cross-culturally (Sharifian 2008; 2011), making these data amenable to questions concerning cross-group variation.

Outside of language research, the task of generating words from predefined semantic categories has been labeled “semantic fluency” or “category instance generation” and is used in neuropsychological studies of word retrieval in patients with brain damage (e.g., Tröster, Salmon, McCullough, & Butters 1989). Similarly, within cognitive linguistics, single-category association tasks are sometimes called “verbal fluency” or “category fluency” tests, where the stimulus can be either semantic or a letter of the alphabet (Bialystok et al. 2010; Gollan,
Montoya, & Werner 2002) and the object of study is the cognitive mechanism of lexical access. The methodology of lexical availability differs in its collection of a broad lexical inventory for the purpose of analyzing vocabulary knowledge and variation on diverse levels.

According to Geeraerts (2010), the study of lexical variation as a sociolinguistic variable has been notoriously challenging and frequently avoided because of methodological difficulties. One has to do with the lack of large enough corpora to capture sufficient instances of the lexical variables. Contrary to the relative ease of collection of sufficient phonological or morphological instances for analysis, the units of analysis in lexical research are larger and thus require a significantly larger corpus to obtain reliable results. Another difficulty is controlling for semantic interpretation in order to measure lexical units in true variation (e.g., Are the terms pants and jeans ever in true variation?). Mitigating this can include the tedious and often inconclusive analysis of the contextual environment of instances, which then negates the intended benefits of big-data analysis.

Although it does not involve naturalistic, textual, or multi-register data (as found in standard corpora and used in variationist studies), the methodology associated with lexical availability is nonetheless a useful tool in the study of the sociolinguistic dimensions of lexical variation. On the contribution of lexical availability research to the study of lexical variation, López Morales (1999) notes that “la orientación fundamentalmente quantitativa que tienen estos estudios sociolingüísticos del léxico disponible los libera de los incovenientes teóricos que afectan a la variación léxica” (p. 25). This is facilitated by an approach based on straightforward categories held constant across participants. Due to the overt nature of the prompts, past research has shown that divergent interpretations or misunderstandings of them are unlikely, and that participant responses, although certainly not uniform, reflect a common construal of the topics
(Bartol Hernández 2010). While this method cannot determine whether words like *pants* and *jeans* are in true variation (Geeraerts 2010), it can provide evidence of the cognitive salience of one relative to the other (in relation to the category of ‘Clothing’) for different speaker profiles. With respect to dialectal variation, Borrego Nieto and Fernández Juncal (2002) outlined a number of advantages of lexical availability relative to methods of traditional dialectology: (1) It is an efficient and quick diagnostic of dialectal characteristics of a specific zone; (2) In relatively little time, it facilitates comparisons between various geographic locations, near or distant; (3) It measures the effect of different semantic categories on the appearance of dialect-specific lexicon.

By way of summary, in addition to its primarily dialectological applications in current research – that is, determining to what extent speakers of distinct geographical regions produce different words, or the same words yet with different frequencies and speed – linguists have used the approach of lexical availability in a number of other fields, including sociolinguistics (to observe social patterns in lexical use), lexicography (to aid in dictionary elaboration), cognitive linguistics (to study the mental mapping of the lexicon), bilingualism (to measure language dominance and contact effects), language pedagogy (to aid in textbook/curriculum design), and ethnolinguistics/cultural linguistics (to observe world perceptions in lexical use).

### 2.5.1 Lexical availability research in Spanish

The use of lexical availability methods in research has been particularly prevalent among scholars of the Spanish language. Many have contributed to a collaborative panhispanic project titled *Proyecto Panhispánico de Léxico Disponible* ‘The Panhispanic Project of Available Lexicon’, headed by Humberto López Morales. The primary goal of this project is the definition of lexically based dialectal profiles through the use of fixed methodological criteria employed across geographical zones and social strata. The bulk of the work to date has documented the
One example of a study using lexical availability data to offer cross-dialectal comparisons is that of Samper Padilla (1999). Given historical migration patterns between the two regions, he compared lexical availability data from the Canarian island *Gran Canaria* and Puerto Rico. He observed that only 13% of the Puerto Rican participants’ words were not found on the Canarian lists, while 20.5% of the Canarian participants’ words were not among the Puerto Rican productions. Present within the vast shared lexicon between the two dialects were items that were unique to both varieties in their own dialectal zones – that is, words unique to Puerto Rico within the Caribbean, and words unique to the Canary Islands within Spain, attesting to dialectal similarities that have withstood time, distance, and other dialectal influences. Similarly, López Morales (1999) observed that the lexicon of Puerto Rican participants overlapped more with the Gran Canarian data (19%) than what it did with data collected in Madrid (14.8%), yet less than what it did with the productions of Dominicans (23.7%).

Lexical availability data based on monolingual Spanish varieties has been used by sociolinguists to determine the level of prestige of specific words (by considering speakers’ socioeconomic status [SES]), to observe in which domains men or women exhibit more lexical diversity, and to test how rural and urban populations differ from one another in terms of lexical robustness or lexical preferences (for a review, see Samper Padilla et al. 2003). In addition, the
SES of a participant has been shown to correlate strongly with the richness of his or her available lexicon.

In this line of work, age has also been found to condition lexical variation. Although studies contributing to the Panhispanic Project as a rule only collect data from final-year high school students (to avoid specialized vocabularies), other work has investigated the available lexicon of the elderly; the results have shown the great extent to which older speakers’ productions can differ from those of young adult speakers within the same community, both due to sociolectal (subcultural) and developmental discontinuities. Borrego (2002, 2004), for instance, found that the lexicon of pre-university students in Salamanca coincided more with that of age-matched speakers in Las Palmas and Puerto Rico than with that of retirees in their same community. This suggests that age may be a stronger predictor of linguistic convergence than geographic proximity, particularly in an era of increasing globalization. In Concepción, Chile, Urrutia (2001), who compared three different age groups, found that the youngest speaker group (aged 15-25) on average produced more words than the older speaker groups (35-40; 65-70), except for in the rather sophisticated fields of problemas del ambiente ‘environmental problems’ and salud y enfermedades ‘health and diseases’. Urrutia also found that the oldest age group’s responses were more homogeneous than those of the other two groups, an outcome Schmitt (2010) attributes to the well-formed lexicons of linguistically advanced speakers. These findings suggest that age-graded quantitative and qualitative lexical variation may be in large part developmental and/or a result of linguistic change across the lifespan.

In light of the goals of the present study, the intergenerational findings in monolingual settings evoke the question of how age differences look in the context of a diaspora community. On the one hand, the G2’s limited use of Spanish in peer communication may diminish
qualitative intergenerational gaps\textsuperscript{14}. On the other hand, in such a context, intergenerational lexical divergence can result from other factors, such as proficiency gaps and mismatched language preferences, or from differences in the conceptual development of the lexicon (e.g., syntagmatically versus paradigmatically motivated responses; circumlocutive or idiosyncratic lexicon; culturally influenced cognition). Selective comparisons of G1 and G2 productions with those of homeland young adult speakers (Valencia & Echeverría 1999) will shed light on these questions.

Psycholinguists have found lexical availability studies useful in determining the mental organization or semantic mapping of the lexicon and the connections that exist among words and between words and word categories. A useful measure in this type of analysis is the “cohesion index” (Echeverría et al. 1987), which calculates the degree of matching in the lexical items produced across participants under a particular semantic theme. In monolingual contexts, results have been rather conclusive with regard to which domains tend to be more “compact” or closed to variation, and which tend to be open and thus more variable across speakers and dialects. For instance, both Alba (1995a), in the Dominican Republic, and Valencia and Echeverría (1999), in Chile, found the categories of \textit{el cuerpo humano} ‘the human body’, \textit{la ropa} ‘clothing’, and \textit{medios de transporte} ‘methods of transportation’ to contain very little variation across speakers, whereas the themes of \textit{profesiones y oficios} ‘professions and occupations’ and \textit{juegos y entretenimientos} ‘games and entertainment’ obtained highly variable responses. In addition, Alba’s (2000) comparative study found a significant relationship between a semantic category’s cohesion index (compactness) and the degree to which different dialects converge in that category.

\textsuperscript{14} Anecdotally, the G2 Chilean participants of this study mused that, to their peers in Chile, their Spanish would likely sound extremely antiquated, their slang outdated.
Given the original motivations for the study of lexical availability, these studies have also been valuable to Spanish language researchers from a pedagogical standpoint. Some Spanish language educators have used (or argued for the use of) native speaker lexical availability results to guide the selection of vocabulary to be included in textbooks for second language learners (Benítez 1994, 1997, 2003; Benítez & Zebrowski 1993; Carcedo González 1999b; Norman 1996). Others have utilized them in tests across age groups or grade levels to measure language development in learners of Spanish as a first language (Echeverría 1991; Samper Hernández 2003) and as a second language (Samper Hernández 2001; Carcedo González 1999a). Carcedo González (1999b) performed an analysis of errors in the available lexicon of second language learners in Finland in order to identify and address common difficulties. In this study he also compared learner vocabulary with that of speakers in Madrid, using the results to propose didactic adjustments in favor of teaching vocabulary that more closely matches usage in a native variety. Lexical availability data has yet to be applied in the curricular development of SHL courses.

Studies of the lexical availability of Spanish speakers have also informed research on language contact and bilingual education. Though far fewer than those based in monolingual locations, several studies have looked at the available lexicon in bilingual communities in and near Spain (e.g., Etxebarria 1996, in Basque Country; Gómez & Gómez 2004, in Valencia; Escoriza 2004, in Gibraltar) and have reported some very interesting findings. Etxebarria 1996), for instance, found that having Euskera as a primary language of instruction at school (as opposed to Spanish) did not contribute to a significantly lower average number of responses by students on a Spanish lexical availability task. However, studies such as these have primarily dealt with bilinguals whose L1 is not an immigrant language (e.g., Euskera, Catalan), and is thus
ascribed a measure of prestige in the region, as well as with speakers who have been formally educated in Spanish and/or who live in zones where Spanish is a language of prestige spoken by the majority (unlike SHS). Offering perspectives on contact, dominance, and thematic knowledge, three studies have examined the lexical availability of Spanish-English bilinguals, including SHS, in the U.S. (Cooper 1971; Moreno-Fernández 2007; Victery 1971 [reviewed in Section 2.4.3]). Nevertheless, further insights can be gained through a more theoretically informed and nuanced use of the lexical availability framework within the social context of Spanish as a low prestige immigrant minority language. In the next section, I outline the social and language factors tested in the present analysis for their relationship with lexical availability.

2.6 Social and Language Experience Factors

Because the acquisitional trajectories of HS are highly heterogeneous, it is important that research on this speaker population consider the effects of different independent variables on the development of the HL. Given its exceptional sensitivity to experience and use (Montrul & Foote 2014), this step is especially critical in studying lexical development. Disparate competencies have been shown to result from a variety of factors, including AoA of the majority language, amount and quality of language input, language attitudes, and residential patterns (see Potowski & Rothman 2011 for a survey of studies reflecting many of these variables in the Spanish-speaking communities of Anglophone societies around the world).

The following social factors, many overlooked or underrepresented in HL research, are considered in this study: exile background; birth order; gender; and parental education. In addition, the relationship between several language factors or measures and lexical output is tested. These include amount of formal HL instruction; Spanish language use; media exposure in the non-majority languages (English and Spanish); English proficiency (discussed in Section
2.4); Spanish proficiency; and proficiency self-ratings in Spanish and English. I will now discuss these factors.

2.6.1 Exile background

Migration classifications have undergone many changes over the last century. Fairchild (1925) classified migrations as invasions, conquests and colonizations. Petersen (1958) later divided them into primitive, forced, free, and massive migrations. David (1970), Fabrega (1969) and others then incorporated the concept of international migration and split the category into two types: voluntary and involuntary. Voluntary migration occurs in response to socioeconomic pressures that indirectly cause someone to choose to migrate. Conversely, the authors describe involuntary migration as resulting from natural disasters, or ethnic, religious, social or political problems. Years later, a new position was posited that questioned the involuntary/voluntary divide in refugee migration (Richmond 1988, 1996). This stance challenged the notion that economically and disaster/socio-politically motivated movements are distinctly determined, and argued instead that an individual’s autonomy is limited in both cases and equally determined by socio-political forces. While a blurred boundary between the two is generally recognized, a distinction continues to be made, given the immediate threat to life, and thus greater urgency of disaster and persecution induced migration.

In the U.S., most Latin American migration is economically motivated, and thus largely voluntary (according to the above criteria). However, in the last 40 years, a sizeable number of Spanish speaking migrants worldwide have resettled due to political unrest in their homelands. The Chilean diaspora is a well-documented example (Wright & Oñate 2005). Although most of the original, or “first wave”, Chilean immigrants in Sweden were political refugees or members of their nuclear and extended families brought through reunification laws, a second wave in the
1980s and early 1990s occurred in response to economic instability (Olsson 2009). Both involuntary and voluntary migrant backgrounds are represented among the participants in this study. This allows for the consideration of exile background as a factor in HL development.

From a sociolinguistic and language maintenance perspective, refugee migration outcomes can be especially varied. Ager (2001) writes that refugees are often afflicted by feelings of inequality and an insecure identity due to low socio-economic status (regardless of former status in country of origin) and/or marginalization. As groups who have a tendency to be labeled and stigmatized (Colic-Peisker & Tilbury 2003), refugees “represent a highly vulnerable community and a group that is at risk of maintaining its heritage language” (Hatoss & Sheely 2009, p. 127). These conditions, coupled with psychological and other trauma associated with the homeland, may lead to fewer positive associations and weaker attachment to their refugee identity, including their HL and culture. However, Colic-Peisker and Tilbury (2003) explain that an opposite pattern of extreme ‘loyalty’ to the original culture and values is also possible because resettlement has been involuntary, sudden, and often perceived as temporary. The mass exile of Cubans in the U.S. following the 1959 revolution was a good example of this. Cuban refugees believed for some time that the new government would not last and that their stay was only temporary, even leaving their belongings in the care of relatives. Over time, the exile beginnings of the Cuban-American population have contributed to romanticized notions of reclaiming the homeland and of the rightful assertion of Cuban identity (Alvarez-Borland 1998; Torres 1999). Thus, the nature of the refugee experience can have important linguistic consequences.

In the case of Chilean refugees in Sweden, Ganuza and King (2005) documented the especially uncertain identities and futures of migrant adolescents who were in legal “limbo” as far as resident/citizenship status, and whose families still contemplated returning to Chile. The
authors describe the difficulties these individuals were experiencing in finding their place in Swedish society and gaining social acceptance. The authors suggest that these circumstances caused individuals to feel more inclined to preserve their Spanish and value their Chilean identities. Conversely, voluntary migrant families may have a more permanent mindset with regard to their residence in Sweden and their need or desire to acculturate. Becker (2013) and Gibbons and Ramirez (2004) suggest that the exile background of Chileans raised in the diaspora favors HL maintenance and development, at least in part due to the complex, politically-oriented discourse in which many become well-versed.

Nevertheless, exile communities have been largely missing from HL research, and from the study of diaspora Spanish in general, despite the majority status of such groups in many settings (Becker 2013; Poyatos Matas & CuatroNochez 2010; Walker 2011), as well as their undoubted ability to enrich our understanding of minority language development and change.

2.6.2 Birth order

The effects of an individual’s birth order have been documented in a number of social and psychological outcomes, including personality (Berthoud 1996), occupational achievement (Conley 2004), intelligence (Kristensen & Bjerkedal 2007), and educational attainment (Booth & Kee 2009). Studies on monolingual language development also point to differences according to birth order (e.g., Pine 1995; Hoff-Ginsberg 1998), generally reporting an advantage for the first-born.

There is also a small body of research treating the effects of birth order on heritage language development in different language situations. These studies also point to the generally stronger proficiency of older siblings, resulting in part from (1) the fact that as children their parents were less likely to have acquired the majority language and (2) the absence of older
siblings providing input in the L2. Jarovinskij (1995) investigated the use of Russian among children of Russian mothers in Hungary, finding that first-borns received more personally directed input in Russian and were more proficient in it than the second- and later-born. In a study among American-Indian families in the U.S., Wong Fillmore (1991) observed a similar pattern: older siblings tended to have a strong command of the home language, while the younger-born were described as less skilled and as having greater English dominance. Pyke (2005) interviewed second generation Asian-American young adults about their siblings and found that “acculturation trajectories” differed across the sibling spectrum, with the oldest ones described as more bound to their ethnic heritage and more likely to speak the parents’ native language fluently. Shin (2002) surveyed a large number of immigrant Korean parents on their children’s language use patterns and also found a significant relationship between birth order and both level of proficiency and the amount of Korean used in parent-child interactions. Using the same methodology (parental questionnaire), Parada (2013) found similar results in a study of 18 Mexican-American families in Chicago. Her results showed that, based on the mothers’ perceptions, birth order was a significant factor in HL abilities, attitudes, and use, particularly between the first/second and third/later-born children.

Stevens and Ishizawa (2007) used population survey data to look at household dynamics in U.S. immigrant families that affect language. With regard to birth order, they found that siblings of the second generation have differing profiles and attributes that influence language development. For instance, older siblings were more likely to be foreign-born and to have spent a larger portion of their lives in the native land. However, they argue that a child’s developmental trajectory can also be affected by these same traits in his/her siblings (i.e., whether they have foreign-born siblings that have lived abroad). Older siblings’ language development may also be
influenced by their more likely designation as family “language brokers” who translate, interpret, and perform other types of linguistic mediation work (Morales & Hanson 2005).

Stockholm’s Latin American population differs from U.S. communities in many ways, including the high levels of Swedish competency among the first generation (thanks in part to widely available government funded Swedish courses), their refugee history, as well as the “minority among minorities” societal status of Latin Americans and of the Spanish language in the migrant landscape. These population characteristics, as well as this study’s analysis of actual participant data (as opposed to parental perceptions) promise to lend greater understanding to the variable of birth order in SHL development.

2.6.3 Gender

Gender-stratified linguistic patterns have been well documented in many areas of sociolinguistic scholarship. In general, women display more conservative behavior with respect to stigmatized variants, yet more innovative behavior with prestige variants and with linguistic change from below the level of conscious awareness (Labov 2001). However, studies of the effects of gender on language maintenance in situations of contact have generated mixed findings.

Klee (1987) investigated the role of gender in combination with SES in a study of Spanish use in the Rio Grande Valley of Texas, finding that Mexican-American men used significantly more Spanish in their daily lives than did their female counterparts. Klee attributes this to the women’s tendency to work in fields that required the use of English, whereas this was not the case for men. She also found that Spanish seemed to be used by males to assert a masculine identity and to build solidarity. English use, on the other hand, seemed more closely associated with a female identity. Solé (1978), in her study among Mexican-American college students, noted that the women were more likely to equate speaking English with greater
socioeconomic opportunities and personal success, and to view Spanish as traditional and socially limiting.

Conversely, in an ethnographic study in a New York Puerto Rican community, Zentella (1997) found that Spanish was used in more female than male domains and that, as a result, girls tended to exhibit greater proficiency in Spanish. Because of the tighter restrictions placed on girls than on boys, who were allowed to leave the block and spend more time in non-Spanish environments, girls had more exposure to and opportunities to use Spanish. In the context of a dual immersion school, Potowski (2004b) also found that female students used an average of 18.4% more Spanish in the classroom than their male classmates, regardless of their L1. She argued that the girls’ greater willingness to “conform to teacher expectations” may have contributed to the disparity (p. 83). In addition, as cited in Section 2.4.3, Victery’s (1971) lexical availability study found that female bilinguals produced more Spanish lexicon than their male counterparts.

The gender outcomes of the socialization of Swedish immigrant youth appear to vary vastly from what Zentella (1997) describes in her work. In fact, Ganuza and King’s (2005) ethnographic study among Chilean-origin youth in Sweden found exactly the opposite: male youth resisted and claimed to encounter more difficulties with cultural and linguistic integration than did their female counterparts (in line Klee 1987 and Solé 1978), citing peer discrimination as the largest barrier. The present study tests whether the gender distinction observed by Ganuza and King (2005) in their interviews bears out in the lexical data analyzed, providing a performance-based view of the gendered linguistic experiences of this speaker population. Although a general comparison of vocabulary size is the focus of the current analysis, the data are able to show how gender plays a role in knowledge across specific semantic domains.
2.6.4 Parental education

Educational attainment is one of the key determinants of socioeconomic status (SES), and is widely accepted as a conditioning factor in language development and use. In contrast with the abundance of sociolinguistic research on the relationship of SES with features of monolingual Spanish, including lexical knowledge (see Section 2.5.1), less is known about its role in HL development. While it is believed that parental views on language and culture influence not only their own practices, but also “their children’s use of and attitudes toward language and culture” (De Houwer 1999; MacGregor-Mendoza 2015, p. 215), how this plays out across minority Spanish community-based scenarios has been inconclusive, with studies so far delivering mixed results. In this study, parental education is considered in the absence of occupation and income data, as the latter two often fail to correlate with educational attainment and the pre-migration SES of migrant populations (Orozco 2007). The parental education of the G2 participants will be tested as a conditioning variable on lexical knowledge.

Concerning the effect of SES on language maintenance, some work (Lambert & Taylor 1996; Sánchez 1983) has found that working class families are more likely to maintain Spanish than those of higher SES because of their limited social mobility and interaction with English speakers in their communities and places of employment. These studies report that, for working class speakers, Spanish use is viewed as important for solidarity building, while middle-class Hispanics regarded English use (at the expense of Spanish) as instrumental to upward mobility. However, in their comparative study of two New York Dominican communities, one working-class and one middle-class, García et al., (1988) found no significant differences in their overall language use patterns. However, although speakers in both communities used lexicon that was
“clearly Dominican”, those belonging to the higher SES community had “more conservative pronunciation” (p. 505).

Conversely, Amastae (1982) found that middle-class Hispanics in Texas placed greater value on Spanish than their lower-class counterparts in the area because, having already achieved a certain level of SES, they were less likely to view Spanish as a hindrance. MacGregor-Mendoza (2015), in her work among Mexican immigrant *profesionistas* (defined as documented, well-educated, English proficient immigrants specifically recruited to fill high skill jobs in the U.S. labor market), found that in their families they “emphasized the development of strong skills in both Spanish and English” and “created an environment where Spanish is valued both privately and publicly” (p. 346). MacGregor-Mendoza argues that it is their emphasis on the value of Spanish in public spaces, as well as the available resources supporting frequent travel to Mexico or that of family members to the U.S. that sets them apart from immigrants of lower socioeconomic standing.

### 2.6.5 Heritage language instruction

To date there is very little classroom or other research that tests the effect of formal instruction on HL proficiency. Two studies to date have tested whether specified instruction on a particular grammatical issue yields short-term results. In Potowski, Jegerski and Morgan-Short (2009), two different methods were used to instruct HL on obligatory past subjunctive contexts with indefinite referents. No statistically significant difference was found by type of instruction, though processing instruction was more effective than traditional instruction. Students did make linguistic gains, however: in terms of interpretation and production, there was moderate improvement, although a grammaticality judgment task showed no statistically significant improvement. Similarly, Montrul and Bowles (2009) were interested in finding out whether
explicit instruction with feedback would help SHS notice and begin to use the preposition “a” when required with animate direct objects. Posttest results revealed significant improvements in both production and intuitions. More generally, it also remains unclear whether amount of seat time in courses designed for SHS is correlated with gains in grammar, vocabulary, or reading/writing. Bylund and Diaz (2012) studied the effects of weekly HL classes on the L1 (HL) proficiency of SHS in Sweden. They found that enrollment in such courses was correlated with better grammaticality judgment and cloze test scores. However, the positive effects were short-term, as proficiency advantages disappeared once students were no longer enrolled.

School-aged SHS in Stockholm do not have access to specialized HL instruction within the daily school timetable. However, “mother tongue” instruction is available by means of a weekly extra-curricular program through ninth grade. There are several restrictions to the provision of mother tongue instruction. According to Hyltenstam (2010, p. 95), “municipalities are responsible for organizing mother tongue instruction in a language if more than four pupils require it and if it is possible to find a suitable teacher.” Students are entitled to this instruction if their language is used on a daily basis in their home and if they have at least a basic level of oral proficiency in the language (Cabau 2014). In 2009, of all students in the nation entitled to this instruction, 54% participated in 146 different languages (Hyltenstam & Österberg 2010). The Swedish National Agency for Education reports that in 2010 only 51.9% of 12,492 eligible Spanish speaking students were enrolled in mother tongue instruction, among the lowest enrollment rates across languages (by comparison, 71% of Somalian and 66% of Arabic speaking students were enrolled). In past decades, when the current study participants were in primary and secondary school, enrollment rates were likely higher due to the peer effects of a larger, more concentrated population of school-aged SHS. Declining Latin American migration
to Sweden has likewise resulted in a decrease in the overall availability of mother tongue instruction in Spanish.

In the present analysis, it is the number of academic years the participants reported having been enrolled in weekly HL instruction that is tested for its relationship with lexical knowledge. Although Bylund and Diaz (2012) found no long-term linguistic benefits, their study did not consider amount of seat time as a continuous variable (only enrollment vs. non-enrollment), nor did they test productive knowledge. The present study will add to the limited research into the linguistic benefits of HL instruction.

### 2.6.6 Proficiency measures and language use/exposure

In order to learn more about the relationship between general HL proficiency and lexical knowledge (Fairclough 2011, 2013a), the results of both an external Spanish proficiency measure and Spanish proficiency self-ratings are considered in relation to the participants’ lexical output. Other factors examined were the participants’ reported Spanish language use and Spanish language media exposure. In order to assess the influence of L3 English competence (as discussed in Section 2.4) on HL knowledge, English external measure performance, English self-reported proficiency, and media exposure to English were also factored into the analysis.

### 2.7 Language Change and the Lexicon

While diachronic work in minority language contexts tends to largely focus on language shift, some work has examined diaspora lexical phenomena strictly in terms of variation and change (e.g., Otheguy & García 1993; Otheguy & Zentella 2012; Nagy 2011; Torres & Potowski 2008; Varra 2013). Otheguy and Stern (2011) argue that many aspects of cross-generational variation in minority Spanish, such as lexical transfer, pronoun expression, and nominal use of the gerund, merely reflect differences, rather than linguistic deficits, resulting from linguistic and cultural
The present study builds on this work by examining patterns of lexical change in relation to language contact in a trilingual context, dialect contact through regular travel to Spain, and culturally induced conceptualization.

2.7.1 Language contact and the lexicon

In situations of language contact, the lexicon appears to be the linguistic level that most readily adopts features of the majority language (Thomason & Kaufman 1988; Thomason 2001). Borrowing, while correlated in some respects with SHS proficiency (Fairclough 2011; Garza 2013, Potowski 2011), must also be understood as a byproduct of cultural contact and as necessary for all speakers, regardless of generation, to achieve expressive adequacy in the context of the host culture (Otheguy & García 1993; Otheguy & Stern 2011). According to Schmid and Köpke (2009), borrowing is a type of cross-linguistic influence that in many cases “constitutes a semantic enrichment of the system, and cannot be taken as evidence for attrition: it is not an indication of previously existing elements no longer being available to the speaker, but of the vocabulary of the speaker (or of the immigrant community) being extended to encompass new concepts and items” (p. 213).

Some research has suggested that lexical borrowing can be quite massive and/or is not necessarily limited to specific semantic domains or word classes (Haspelmath 2008; Thomason & Kaufman 1988, Torres & Potowski 2008a). Other work on heritage Spanish has shown that the degree of majority-language influence it exhibits can in fact be rather insignificant. For example, in a New York Cuban Spanish spoken corpus, Otheguy, García and Fernández (1989) found that only 6.9% of the second-generation 5,724 words analyzed was English-influenced. In an early analysis of their Chicago Spanish corpus, Torres and Potowski (2008b) found that among a total of 117,024 words produced during hour-long sociolinguistic Spanish interviews,
only 5.1% were in English. Even in receptive bilingual writing samples, (Fairclough & Belpoliti, 2015), a relatively low percentage (4.95%) of lexical transfer was found.

Similarly, we can recall from Section 2.4.3 that among the thematic word lists generated by SHS in Chicago, only 15% of the more than 3,000 lexical items produced were Anglicisms (Moreno-Fernández 2007). When he instead considered only the group’s first 20 items in each thematic category (i.e., the most readily mentioned and least idiosyncratic words of the group’s productions), the percentage decreased to 6.5% (of 400 total lexical items). Additionally, the author noted that a significant portion of the Anglicisms were produced superfluously, as they were offered in doublets with their Spanish counterparts. The remainder of instances he attributes to lexical gaps, given their appearance in categories less related to the speakers’ daily activities as well as their correspondence to more technical or low frequency lexicon. Two of the categories in which he recorded a high number of Anglicisms were: juegos y distracciones ‘games and entertainment’; and calefacción, iluminación, y medios de airear un recinto ‘heating, lighting, and methods of ventilating a space’. Alternatively, in the categories of el cuerpo humano ‘the human body’ or animales ‘animals’, no Anglicisms were produced. The absence of borrowings in the more routine areas of the lexicon was also noted by Gamboa González’s (2003) (see Section 2.4.3).

These findings align with other research showing that loanwords tend to appear last in basic lexical domains, where “basic” refers to words used to describe the most typical and essential human activities or needs (Hock & Joseph 2002). Basic vocabulary is more “entrenched” in a language, and is therefore believed to be more resistant to change (Nagy 2011). Therefore, when borrowing reaches the most basic domains, it is indicative of an advanced stage of language contact (and, perhaps, at that point, language shift). Similarly,
Myers-Scotton (2002) draws a contrast between *cultural* and *core* borrowings, which she explains come about differently. Cultural borrowings are words labeling new objects (e.g., *espresso*) or concepts (e.g., *zeitgeist*) and they readily emerge when associated with groups of power. Core borrowings are words that appear to duplicate already existing words in the receptor language (e.g., ‘OK’ in German, replacing *gut*, or *einverstanden*; ‘gym’ in Spanish, replacing *gimnasio*) and often begin as code-switched utterances.

This pattern is supported in research on Spanish in the U.S. in which, among common English borrowings, a predominance of cultural and conceptual terms are found. These denote novel (or appreciably different) objects or concepts encountered post-migrationally. What qualifies as a new concept, however, may include quite subtle cross-cultural differences. Otheguy and García (1993) argue that even when there is a seemingly equivalent word in the language of origin, its referent may differ in significant ways. As mentioned earlier, they found, for example, that Spanish speakers in New York used both the Spanish *edificio* (to reference types of buildings in Latin America) and the English */bildung* to describe structures common to their U.S. American experience. Thus, it appears that many borrowings are known in addition to and occur in contrastive distribution with related homeland or general Spanish terms (Otheguy & Stern 2011).

Another distinction made in analyses of lexical transfer is that of nonce (spontaneous) versus established borrowings used frequently in the community, the former reportedly more common among second- and later-generation speakers (Lipski 2005; Poplack et al. 1988). Sánchez Muñoz (2010) designates nonce borrowings in heritage Spanish as ‘lexical creations’, describing them as infrequent in bilingual Spanish and often belonging to one’s idiolect. In his analysis of the Anglicisms produced by SHS in Chicago, Moreno-Fernández (2007) observed
both; in most cases, the borrowings were produced by two or more participants (e.g., *yarda* ‘yard’; *troca* ‘truck’), yet others appeared to be spontaneous in nature (e.g., *forka* ‘fork’; *mojar plantas* ‘water plants’).

The terms ‘lexical transfer’ and ‘borrowing’ denote a number of contact phenomena. Given the nature of the lexical availability task, this study examines lexical borrowings that may be identified in a non-discursive written word association exercise. Below, I provide a typology of borrowings in which are outlined the specific contact phenomena included in the analysis. It is adapted from Mendieta’s (1999) *tipología del préstamo léxico*. Mendieta defines *préstamo léxico* as “la incorporación de una unidad lexi*a L2 (o de un compuesto que funcione como unidad lexic*a) en un contexto L1.” In the case of the current study, L1 refers to Spanish while L2 corresponds to either Swedish or English.

**TABLE 2-4**
BORROWING TYPOLOGY

<table>
<thead>
<tr>
<th>Mode of Transfer</th>
<th>Borrowing Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporation</td>
<td>Pure borrowings</td>
<td>marqueta (<em>mercado</em>) “market”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>averaje (<em>promedio</em>) “average”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>troca (<em>camioneta</em>) “truck”</td>
</tr>
<tr>
<td>Reproduction</td>
<td>Phrasal calques</td>
<td>tener un buen tiempo (<em>divertirse</em>) “to have fun”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>investir (<em>invertir</em>) “to invest”</td>
</tr>
<tr>
<td>Incorporation + Reproduction</td>
<td>Loanblends</td>
<td>calendador (<em>calendario</em>) “calendar”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tripear (<em>hacer tropezar</em>) “to trip (somebody)”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pantimedia (<em>medias</em>) “pantyhose”</td>
</tr>
<tr>
<td>--</td>
<td>Phonological Intrusions</td>
<td>mlútica (<em>música</em>) “music”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>quémica (<em>química</em>) “chemistry”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>narcáticos (<em>narcóticos</em>) “narcotics”</td>
</tr>
</tbody>
</table>
Following Haugen (1950, 1967), Mendieta (1999) broadly classifies borrowings according to the manner in which they are transferred into the L1: incorporation or reproduction. The first refers to the incorporation of L2 morphemes into the L1, which may reflect varying degrees of phonological adaptation. This mode of transfer comprises the borrowing category of pure borrowings. The second mode of transfer, reproduction, involves the mapping of L2 meanings or structures onto L1 morphemes. Phrasal calques and semantic extensions correspond to this category. While pure borrowings are highly frequent and widely accepted in both bilingual and monolingual settings, phrasal calques and semantic extensions are primarily limited to situations of intense contact (i.e., high levels of bilingualism). Both the incorporation and reproduction modes of transfer are involved in the category of loanblends, which contain L1 and L2 morphemes. These can include L2 borrowings that exhibit L1 morphological adaptation (verbal or plural morphemes, for example), or compounds of unbound L1 and L2 morphemes (e.g., pantimedia). All of the preceding types of lexical transfer can be additionally classified in terms of their status as spontaneous or established (shared) borrowings in a given community. A final category of Mendieta’s typology deals with what Teschner (1972) originally labeled phonological intrusions. These include words that exhibit L2 influence in the pronunciation or written form of an L1 cognate (e.g., miúsica), usually as an effect of low literacy in the HL.

The L1 Spanish of both G1 and G2 speakers in the Chilean-Swedish community has experienced intense contact with Swedish. Thanks to extensive governmental provisions, adult migrants to Sweden are given ample resources to acquire the majority language, which generally begins soon after arrival. In addition, as discussed in Section 2.3, the G2, and even some G1 speakers, typically develops considerable fluency in English. This occurs through obligatory formal instruction as well as the ubiquity of English in Swedish media and popular culture.
(Sharp 2001; Pennycook 2014). Evidence of English influence was thus expected in the lexical productions as well. Based on the forgoing evidence, it was expected that borrowings produced by the Chilean-Swedish participants of this study would be relatively few in number, and have a greater likelihood of appearing in (1) domains calling for less routine and more abstract vocabulary and, relatedly, (2) in domains in which concepts/objects of the host culture figure prominently (e.g., \textit{la política}). It was also expected that G2 speakers would be mostly responsible for any phrasal calques, semantic extensions, or phonological intrusions in the data.

In sum, of principal interest in the inter-group analysis was (1) the density of Swedish versus English borrowings in each tested domain, and (2) how the borrowings varied on the following dimensions: borrowing type; nonce versus established.

\textbf{Task effects}

The contact phenomenon of code-switching is conspicuously missing from the typology presented in Table 2-4. This is because the non-conversational nature of the lexical availability task as well as the explicit instructions for participants to generate items in Spanish renders it irrelevant to the current analysis. Therefore, any single or multiple-word lexical item bearing influence from Swedish or English was analyzed in accordance with the borrowing typology in Table 2-4.

An additional task effect warranting mention involves the idea that written tasks, perceived as inherently formal, may tend to elicit fewer borrowings than natural speech. This perception could also mean that they are simply produced less readily, appearing in the latter portion of participants’ word lists (Hernández-Munoz 2009). It is possible that these effects pertain mainly to the G1 responses, given that HS tend to practice less style-shifting and often may not recognize borrowings as such (Sánchez Muñoz 2010; Valdés & Geoffrion-Vinci 1998);
different studies have documented the blurred boundaries between oral and formal lexicon exhibited in HS writing (Colombi 1997; Fairclough 2006; Schwartz 2003). As for spontaneous lexical transfer, G1 and G2 patterns may differ substantially. Sánchez Muñoz (2010), for example, found that greater task formality actually correlated with an increased oral production of lexical creations (e.g., endurar “to endure”) by SHS. Moreno-Fernández (2007), as well, observed a number of spontaneous Anglicisms (e.g., mojar plantas ‘water plants’) in written HS responses to Spanish prompts. These reports suggest that the spontaneous production of borrowings by HS in certain contexts may actually reflect an attempt to use sophisticated language by (perhaps unwittingly) drawing on stylistic knowledge in the dominant language.

Regarding the limitations of the written nature of the lexical availability task, Hernández Muñoz (2009) suggests a silver lining: the fact that the written mode does not seem to favor the production of highly regional vocabulary (including borrowings) lends special status to the those items actually mentioned, revealing their integrated status in the speakers’ lexical inventories. However, an important consideration is whether the widespread use of informal written language via social media may be contributing a weakening of the written–formal association.

### 2.7.2 Dialect contact

Research has shown that, like acquiring a second language, speakers can acquire features of a different dialect of their first language (Siegel 2010). As Schmidt (2015) points out, although native-like acquisition of a second language (e.g., Hyltenstam & Abrahamsson 2000; Johnson & Newport 1989; Krashen 2006) or second dialect (e.g., Labov 1970; Chambers 1992) is sensitive to maturational constraints, language is subject to change throughout a speaker’s life. As we have seen, the lexicon is especially disposed to ongoing change.
Relocation to a new dialect region is one way dialect contact is occasioned, and can lead to a speaker's adoption of new linguistic features. Several studies have traced the acquisition of dialectal features by children who have moved to a new dialect area (Chambers 1992; Kerswill 1996; Payne 1980; Starks & Bayard 2002; Tagliamonte & Molfenter 2007; Trudgill 1986). These studies found that the children all acquired some of the phonetic, phonological, and lexical features of the new dialect, but complete dialect acquisition appeared to be constrained by the complexity of features as well as the age at which contact began.

The speech of adults may also undergo change as a result of moving to a new dialect region (e.g., Munro, Derwing & Flege 1999; Evans & Iverson 2007) or as a result of dialect contact in a minority language context (Potowski & Torres, forthcoming; Zentella 1990). In a study of adult Canadians who had relocated to Alabama, Munro et al. (1999) used ratings provided by speakers from the native and new dialect regions to determine the degree of dialect change. Raters of both dialects perceived an intermediate degree of the regional accent in the speech of the adult Canadian transplants, which indicated partial acquisition of some of the phonetic features of the second dialect. Likewise, Evans and Iverson (2007) found that speakers of northern British English, a nonstandard dialect, in contact at university with the prestigious Standard Southern British English (SSBE) were rated as sounding more "southern".

Spanish dialect contact among adult speakers in a minority context has also been addressed in a handful of studies. With respect to the lexicon, two studies have examined other-dialect familiarity among U.S. Latinos of different origins. Zentella (1990) tested the factors conditioning degree of other-dialect familiarity among 194 New York Latinos of Colombian, Puerto Rican, Dominican, and Cuban heritage. Although the speakers varied in terms of age, gender, education, years in the U.S., and English and Spanish proficiency, generational
information was not provided. Zentella’s experiment consisted of showing pictures of 25 different objects, and asking the participants how they would normally refer to the items, as well as if they were familiar with any additional terms used by other Spanish speaking groups. The objects were selected because of their presence in daily life and because there were two or more different words used to refer to them among the dialects in question. Zentella found that socioeconomic status, level of education, and race/phenotype played a greater role in the directionality of dialect influence than did numeric superiority in the area. Despite the greater number of Puerto Rican and Dominican Spanish speakers, the forms used by the more educated and lighter skinned Colombians and Cubans were the most common other-dialect items known among the participants. In the cases of objects for which there was significant lexical diversity across dialects (e.g., bus; birthday cake), she found that English variants emerged as communicative neutralizers. Zentella concluded that linguistic change and diffusion are shaped by “a number of social and economic realities” (1990, 1097).

Potowski and Torres (forthcoming) used similar methods in their study of contact between Mexican and Puerto Rican Spanish varieties in Chicago. They showed images of 10 objects (orange, pacifier, bus, earrings, banana, etc.) to 37 Mexicans and 39 Puerto Ricans representing three speaker generations. The experiment was also administered to control groups of 15 homeland Mexicans and 15 homeland Puerto Ricans. The authors predicted that, due to the greater number of Mexicans in Chicago, Puerto Ricans would be more familiar with Mexican vocabulary than the reverse. Another prediction was that the G1 speakers would know more other-dialect forms than the G2, who would, in turn, know more than the G3. Indeed, the Puerto Rican participants did demonstrate more familiarity with Mexican vocabulary than the Mexicans did with Puerto Rican words. Although a similar result was found for the homeland groups, they
did not exhibit as much exogroup vocabulary knowledge as their counterparts in Chicago. With respect to generation, the results only partly matched the authors' prediction; while the G3 did in fact know the least exogroup vocabulary, the G2 surpassed the G1 in their level of familiarity.

Via the mediums of social contact, media, and travel, exposure to non-local dialects may also occur. Studies on English varieties in contact through media exposure (Scott & Cutler 1984; Evans & Iverson 2004; Stuart-Smith et al. 2013) have generally found that it induces no measurable change in the language of the consumer. With respect to Spanish, little is known about the linguistic effects of non-local dialect exposure. Schmidt (2015) considered this question in studying the role of exposure to /s/-weakening dialects in the perceptual categorization of syllable-final /s/-aspiration. The study was conducted among young adult Bogotanos in Colombia, whose dialect does not include this feature. She found that “changes in perceptual norms of dialectal sounds may occur even in the absence of an intense immersion experience (i.e., moving to a second dialect region)” (p. 100). Specifically, Schmidt’s (2015) results revealed that the presence in the participants’ social networks of speakers whose dialects include this feature played a significant role in their identification of the regional variant. However, the participants’ exposure to /s/-weakening through media and short-term travel (with an average stay of a month) did not contribute to perceptual change.

According to Schmidt (2015), her findings underscore the special role of “live social interaction” (p. 113) in dialect change, echoing Labov’s (1994) emphasis on the importance of “face-to-face interaction with peers” for language change to occur. In essence, the quality and quantity of dialect exposure determined whether speakers’ perceptions were altered. Schmidt (2015) concedes, however, that the results regarding the effects of media are preliminary, and that “future work is needed to further understand the short- and long-term linguistic effects of
exposure to dialectal speech through this medium (music, film, television, etc.)” (p. 114).

Furthermore, with regard to short-term travel, she suggests that taking into account the recency of travel and the amount of interaction between the travelers and local dialect community could provide more nuanced results. In addition, I would add that it may be important to consider whether there has been repeated short-term travel to a single destination, given that this would increase the likelihood of interaction with local dialect speakers as well as their level of integration into the travelers’ social networks.

The goal of this portion of the current study was to investigate how forms of non-local dialect contact (i.e., travel, media) might influence a minority variety of Spanish. As the majority Spanish speaking group in Sweden, Chilean-Swedes likely make few accommodations to the other Spanish varieties represented in their communities (Potowski & Torres forthcoming; but see Zentella 1990). However, they regularly experience contact with Peninsular varieties due primarily to Southern Spain as a frequent vacation destination, but also to the widely available Spanish television networks in Swedish media (transmitted in Castilian Spanish). Questionnaire data revealed that 39 of the 53 (74%) study participants had traveled to Spain at least once. The majority (n = 33) of the 39 had visited on multiple occasions, having spent an estimated lifetime average of 100.6 total days in that country. Several of the participants even owned condos or had time-share arrangements, which they (informally) reported had contributed to the development of long-term friendships and association with Spanish locals. Though the participants' ages at the time of travel were not requested, informal conversations with the G2 participants left the impression that many of them had spent time in Spain as minors. Given the greater propensity of children to acquire new dialectal forms, it was possible that the G2 speakers' lexicons would exhibit greater dialect contact effects.
Considering the minority status of Chilean vacationers in Spain and that of Chilean Spanish in Europe more broadly, the many factors favoring the general attribution of prestige to Peninsular Spanish (Villa 2002), as well as the high probability of interaction with target community members, it was conceivable that the participants’ vocabulary would exhibit transfer from the varieties with which they have experienced repeated contact. Thus, it was of interest to determine whether lexical items unique to Peninsular Spanish were present in the participant productions, whether there were generational differences, and how prominent these items were in the word lists. Because the lexicon tends to be first acquired in second dialect acquisition (Siegel 2011), it seemed probable that some traces of contact would be observed in the data.

2.7.3 Cultural linguistics and the lexicon

A final aspect of cross-generational language change examined in this study is that of conceptual differences (Jarvis & Pavlenko 2008; Polzenhagen & Xia 2014; Sharifian 2011), which can be discerned from availability rankings (Carcedo González 2000; Ferreira & Echeverría 2010; Mackey 1971; Hernández Muñoz et al. 2006). In lexical availability research, availability rankings are arrived at mathematically by considering both the frequency (number of tokens) and readiness of mention of lexical items across speakers of a given group. Data are codified and entered into a database (www.dispolex.com) where a formula is applied that generates an availability ranking or “index” for each lexical item (type) produced by a participant group for a given prompt.

This kind of analysis allows for conceptual differences to be observed across speaker groups, who may show unique patterns in the ease of mention of certain items, or in the presence/absence of others, in relation to a given prompt. Such patterns can reflect sociocultural aspects of speakers’ collective life experiences, and provide insight into the way particular topics
are conceptualized. In other words, the degree of availability of a lexical item within the context of a specific semantic domain may be viewed as a function of its prototypical value for the speaker group in question.

The organization of the mental lexicon is understood as bound up with categories of experience, rather than in the strict format of, say, a dictionary. Research shows that word/concept *familiarity* and *typicality* are two factors that heavily influence which lexical items will appear first in the mind when elicited through a thematic stimulus (Hernández Muñoz et al. 2006, p. 746):

(1) *concept familiarity* - words denoting very familiar objects or concepts are more available for use than words denoting less familiar objects or concepts;

(2) *typicality* - words that are judged to be highly typical exemplars of a chosen category are generated by more people earlier in their lists

In the introduction, I referenced the utility of Rosch’s (1978) theories of categorization and prototype gradation as a framework to explore cross-linguistic conceptual discontinuities in different thematic areas (e.g., colors, emotions, lying). According to Polzenhagen and Xia (2014, p. 19), this line of work is “devoted to research that highlights cultural dimensions of prototypicality and their linguistic manifestations”, motivated by the rationale that “prototypicality is context sensitive and context dependent”; accordingly, it appears that sociocultural context both “influences and establishes prototypicality.”

Within the discipline that he terms *Cultural Linguistics*, Sharifian’s (2011) theories of conceptualizations at the cultural level of cognition also provide an important backdrop for the question of how distinct socializations and cultural experiences can explain inter-generational conceptual differences. This field, sometimes also referenced as *Ethnolinguistics*, is concerned with the ways world perceptions and conceptualizations shape language. Sharifian’s model of
cultural conceptualization, which incorporates the cognitive notions of schemas and categories, illustrates how conceptualizations are “distributed across the minds constituting a social group” (p. 3). In traditional connectionist models of cognition, schemas and categories are understood not as items in the mind, but as patterns that emerge in the organization of knowledge units. While such patterns are considered properties of an individual’s cognition, Sharifian (2011, p. 3) argues that they also operate at a cultural level:

Although it is admitted that the locus of conceptualisation may be the individual, a large proportion of conceptualisations are ultimately ‘spread’ across a cultural group. In other words, although conceptualisations can be initiated in individuals’ cognition, they may well emerge as cultural cognitions.

However, Sharifian also contends that group member interactions can produce emergent conceptualizations that may not be reduced to conceptualizations of the individual minds. Central to Sharifian’s theory of cultural conceptualization is the idea that cultural groups are not formed solely on the basis of shared physical space but also on individuals’ participation in a shared conceptual world. In fact, “the degree to which individuals can participate in a group’s conceptualized sphere determines their membership in the group” (p. 5).

In addition to the study of cross-linguistic conceptual variation, other research has examined how cultural and environmental differences affect the prototype structures of different varieties of a single language. Schmid et al. 2008, for example, studied the idea of “culturally blended concepts” in the context of “New Englishes” in West Africa. Under the assumption that these varieties bear conceptual influences from both English speaking societies and local languages and cultures, the authors compared the attribute prototypicality of common lexical items such as airplane or wheelchair for three speaker groups: native speakers of U.S. American English; speakers of Nigerian (Huasa) English; and native speakers of Huasa. They found that
the Nigerian English speakers’ conceptions displayed a variable mixture of components from their two cultural affiliations.

Similarly, discourse studies have documented how miscommunications occur between Aboriginal and mainstream Australian English speakers due to conceptual differences, sometimes even leading to disadvantages in places like the courtroom or classroom (Christie & Harris 1985; Eades 2000). Sharifian (2011) describes how in these scenarios even “everyday words such as ‘home’ and ‘family’ may evoke contrasting cultural conceptualisations” (p. 79). In addition, the perceptions of concepts such as ‘sing’ and ‘smoke’ vary significantly between these speaker groups (containing especially spiritual connotations for Aboriginals), though this is not always explicit in interaction. Sharifian (2011) points out that while interlocutors may be able to get a rough idea of the significance behind words like these for Aboriginals, understanding the events to which they are tied and “the emotions associated with [them] would require being equipped with the conceptualisations which are rooted in the particular view of the world” (p. 82).

Other studies have used lexical measures to look at how native speakers versus L2 learners compare in their conceptual organization of different domains. In second language acquisition, speakers’ L1 conceptualizations have been shown to directly influence semantic mappings in the L2. Jarvis and Pavlenko (2008) refer to this as semantic or conceptual transfer. An example from Spanish is Carcedo (2000), who used a lexical availability task to compare conceptual differences between native speakers of six different Spanish-speaking regions and L2 learners of Spanish in Finland. He found that unlike any of the native respondents, under the topic of *el campo* ‘the countryside’, the Finnish learners consistently placed the words *bosque* ‘forest’ and *lago* ‘lake’ at the top of their lists. This conceptual contrast
reflected an upbringing in Finland where forests and lakes are prototypical elements of the outdoors. Similarly, in her study of Slovene L2 learners of English and Spanish, Šifrar Kalan (2014) noted the prominence of the word/concept “green” in the students’ lexical availability lists. This research supports the idea that the environment (physical, cultural, educational, etc.) of one’s upbringing strongly influences conceptual organization, which is in turn reflected linguistically.

In the present study, I ask whether there is meaningful variation between the speaker groups in conceptualizations of the six domains analyzed. Specifically, I study the domain set members of high prototype status (i.e., high availability ranking) for indications of conceptual divergence.

2.8 Conclusion

To sum up, this study offers a comparative analysis of the lexical repertoires of first and second generation ethnic Chileans in Stockholm, Sweden, where Swedish and English govern all public life. The objective is to supply detail on the maintenance of Spanish in this community, including how particular socio-demographic variables condition lexical aspects of heritage language development. Importantly, the project builds on previous investigations of SHS lexicon by centering the analysis on productive knowledge contextualized within semantic domains falling along a basic-to-abstract continuum.

In addition, the study is concerned with dialectological aspects of this diaspora variety. Extending beyond a deficit approach, it aims to identify and evaluate some of the qualitative lexical shifts that the community’s variety is experiencing. This assessment of inter-generational qualitative lexical change includes analyses of lexical transfer resulting from language and dialect contact, as well as of culturally induced conceptual shifts represented in the mental
lexicon. In short, the multi-categorical data permit both quantitative and qualitative perspectives of the lexical systems of an understudied trilingual contact variety of Spanish.
3. METHODOLOGY

3.1 Introduction

In this chapter I discuss the participant characteristics and recruiting steps followed, the data collection tasks, and the coding, tools and statistical measures employed in the analysis of the data. Motivated by the review of literature presented in Chapter 2, I restate my research questions (RQs) below:

1. What are the quantitative characteristics of SHS lexicon, as measured in a lexical availability task?
   a. Is there a significant numerical difference between SHS (G2) and first generation (G1) speakers in the words produced for each of 21 semantic domains? What role does domain complexity play?
   b. What is the relationship between SHS lexical knowledge (number of words produced) and the following social variables: gender; birth order; parental education; and exile background (political vs. other)?
   c. What is the relationship between SHS lexical knowledge (number of words produced) and the following language variables: amount of HL instruction; performance on a general Spanish proficiency measure; performance on an English proficiency measure; self-ratings of Spanish and English proficiency; media exposure to these languages; and Spanish language use?

2. What are the qualitative lexical characteristics of the participant productions in six semantic domains? Specifically, how do the following reflect inter-generational change?
   a. Response compatibility rates and characteristics of divergent lexicon
b. The density and types of Swedish and English borrowings

c. The density and types of Peninsular Spanish regionalisms

d. Category conceptualizations

3.2 Participants

The data collection was realized in the metropolitan area of Stockholm, Sweden, where adequate access to the Chilean-Swedish population was expected. Given the minority status and scattered residential patterns of ethnic Chilean individuals in Stockholm, recruiting a sufficient number of participants presented somewhat of a challenge. Lexical availability research in monolingual settings, due to its quantitative interests, has traditionally relied on a large participant pool; for each dialectal area, the norm has been to survey between 80 and 400 individuals, depending on the population size studied. This is typically accomplished by visiting high school or first-year university classrooms and gathering data one class at a time. However, following Moreno-Fernández (2007), whose analysis included data from 48 bilingual heritage speakers in the U.S., I realistically aimed to recruit at least 60 total participants. In the end, a total of 33 G1 and 30 G2 participants were recruited. However, incomplete data resulted in the elimination of three G1 participants. In addition, the data provided by seven participants in the G2 group were excluded due to their background as G1.5 speakers; because their ages of arrival and transmigration patterns exhibited wide variation, and they were too few in number to form a third comparison group, the decision was made to exclude their data from analysis. The final participant totals rested at 23 G2 and 30 G1 (total n = 53).

The participants were recruited through family, friendship and community networks, as well as on local university campuses. Because the tasks were compatible with completion in a

\[15\] According to the criteria set forth in Silva-Corvalán (1994), G1.5 speakers include those who arrived to the U.S. between the ages of 6 and 12.
group setting, efforts were made whenever possible to organize group sessions. These meetings were held at participants’ homes and at local facilities such as libraries and churches. All individuals were required to sign a consent form prior to participation and were monetarily compensated (roughly $25). The G1 age range was 44-72 (average 54.8 years) and the G2 age range was 18-42 (average 25 years). The G1 group consisted of 14 females and 16 males, while the G2 group was composed of 15 females and 8 males. All but five of the G2 participants had been born in Sweden, and those five all arrived before the age of three. The mean number of years the G1 participants had lived in Sweden was 25.3, or 46% of their lives.

3.3 Data Collection Procedure

The data collection took place in a single session lasting approximately 90 minutes. A background and language use questionnaire occupied the first 15 minutes. Next, a Spanish proficiency measure was allotted 20 minutes, and an English proficiency measure another 10 minutes. A series of word association (lexical availability) tasks occupied the remaining 45 minutes.

3.3.1 Questionnaire

The questionnaire is included in Appendix A. The background portion requested the participants’ gender, age, educational background, parents’ educational background, neighborhoods of residence, place of birth, birth order, age of arrival in Sweden (if born abroad), along with other details such as the primary reason for one’s (or one’s family’s) migration, the likelihood of the family’s returning permanently to Chile, travel patterns to Chile and other Spanish-speaking countries, and social network details. Although income and occupation were also requested, the response rate was poor; these are admittedly sensitive and complex issues, especially since many of the exile background participants had involuntarily abandoned their professions and university
studies in Chile with little promise of resuming them in Sweden. Table 3-1 charts the factors considered in RQ1b.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Predictions based on previous research</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total lexical output</td>
<td>participants’ total word production across 21 thematic domains (domains listed in Table 3-4)</td>
<td>Highly heterogeneous between-speaker production rates</td>
</tr>
<tr>
<td><strong>Independent sociological variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>female, male</td>
<td>Lexical proficiency advantage for male speakers</td>
</tr>
<tr>
<td>Birth order</td>
<td>oldest, middle, youngest</td>
<td>The higher the birth order, the more robust the HL lexicon. Significant differences after the second-born.</td>
</tr>
<tr>
<td>Parental education</td>
<td>average of mother and father. based on a five point scale: 1 - primary education, 5 - advanced degree</td>
<td>SHS whose parents have gone to university will display stronger lexical knowledge.</td>
</tr>
<tr>
<td>Exile background</td>
<td>politically motivated migration; other</td>
<td>SHS from an exile background will display stronger lexical knowledge.</td>
</tr>
</tbody>
</table>

A second section of the questionnaire requested language related information. This included the number of school years G2 participants had enrolled in hemspråk, or “mother tongue”, instruction, as well as integrated self-ratings of literacy, writing, speaking, and comprehension abilities in each of their languages (Spanish, Swedish, English). In addition, all participants completed an in-depth Spanish language use survey. They were also asked to rate how often they were exposed to Spanish and English through media (music, TV, reading). To complement the self-reported data, participants also completed proficiency measures in Spanish (G1 and G2) and in English (G2). A description of these is provided ahead. It was expected that these data would permit a nuanced account of the past and present functions and relative
importance of the three languages in their lives. Table 3-2 summarizes the language related variables considered in the factorial analysis.

### TABLE 3-2
SUMMARY OF LANGUAGE VARIABLES

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Predictions based on previous research</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total lexical output</strong></td>
<td>Participants’ total word production across 21 thematic categories</td>
<td>Highly heterogeneous between-speaker production rates</td>
</tr>
<tr>
<td><strong>Independent language variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External proficiency measure - Spanish</td>
<td>% right</td>
<td>Positively correlated with HL lexical output</td>
</tr>
<tr>
<td>External proficiency measure - English</td>
<td>% right</td>
<td>Positively correlated with HL lexical output, possibly due to beneficial cross-linguistic connections,</td>
</tr>
<tr>
<td>HL instruction</td>
<td>Number of academic years</td>
<td>Positively correlated with lexical output.</td>
</tr>
<tr>
<td>Spanish language use</td>
<td>30 questions related to daily language use. Answers on a five-point scale: 1 – almost never/never, 5 almost always/always. Average calculated.</td>
<td>Positively correlated with HL lexical output.</td>
</tr>
<tr>
<td>Proficiency self-rating - Spanish</td>
<td>Proficiency in reading, writing, comprehension, and speaking rated on a five-point scale: 1 – poor, 4 – very good. Average calculated.</td>
<td>Positively correlated with HL lexical output.</td>
</tr>
<tr>
<td>Proficiency self-rating - English</td>
<td>Proficiency in reading, writing, comprehension, and speaking rated on a five-point scale. Average calculated.</td>
<td>Positively correlated with HL lexical output, possibly due to beneficial cross-linguistic connections,</td>
</tr>
<tr>
<td>Media exposure Spanish</td>
<td>Print, TV, music exposure. Rated on a five-point scale, 1 – never/ almost never, 5 – always/almost always. Average calculated.</td>
<td>Positively correlated with HL lexical output.</td>
</tr>
<tr>
<td>Media exposure English</td>
<td>Print, TV, music exposure. Rated on a five-point scale, 1 – never/ almost never, 5 – almost always/always. Average calculated.</td>
<td>Positively correlated with HL lexical output.</td>
</tr>
</tbody>
</table>
A Spanish proficiency measure external to the lexical task was considered important to understanding how lexical knowledge relates to other aspects of proficiency (Fairclough 2011, Polinsky 2006). This step represents an improvement over Moreno-Fernández (2007), who judged proficiency according to SHS course level. All participants were administered the multiple choice cloze and vocabulary portions of the Diplomas of Spanish as a Foreign Language (DELE) test (see Appendix B), and their scores (total number correct) were recorded. The successful completion of a cloze test involves drawing on a number of competencies, including contextual, morphosyntactic and lexical knowledge. In Montrul’s work (2004, 2007, among others), this instrument has generally, though not consistently (Montrul & Perpiñán 2011) been effective in distinguishing fine-grained differences among speakers, including SHS of differing profiles. Other measures were considered, such as the Elicited Imitation Test (EIT) and the Simulated Oral Proficiency Interview (SOPI), but these take much longer to administer. Given the time constraints accompanying data collection abroad and in a minority community, the DELE was ultimately adopted for use in this study, complemented by the participant self-ratings.

In order to explore the question of how L3 English proficiency relates to HL development, the participants also completed a brief English proficiency measure. This was administered to all G2 participants, as well as to two G1 participants (who opted to take it). The measure used was a modified version of the Integrated Levels Productive Vocabulary Levels Test available online (www.lextutor.ca/tests/levels/productive/). It is described as a 30-item controlled productive ability test based on words belonging to three frequency levels (Laufer & Nation 1999): 2000, 3000, and 5000. The format was simple; discretely presented sentences each contained a missing word for which the first three letters were already provided. The instructions
invited the learner to complete the words without undue regard for correct spelling. An example was provided. Table 3-3 shows the distribution of items and accompanying examples.

<table>
<thead>
<tr>
<th>Frequency level</th>
<th>n</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>8</td>
<td>I'm glad we had this opp__________ to talk.</td>
</tr>
<tr>
<td>3000</td>
<td>11</td>
<td>He has a successful car__________ as a lawyer.</td>
</tr>
<tr>
<td>5000</td>
<td>11</td>
<td>A bird perched at the window led__________</td>
</tr>
</tbody>
</table>

3.3.2 The lexical availability task

In lexical availability research, the material is obtained by means of associated tests which present thematic categories for which participants contribute lexical items considered related (Etxebarria 2014). Timed, unaided responses are an integral part of the methodology, making a proctored environment essential; that is, reliable results require that the data be a true reflection of each individual’s independently produced lexicon. Because the aim is to uncover the lexicon that speakers are most likely to use in real life communicative contexts, a two-minute time limit per category ensures the production of the most readily available lexical items.

Given the time constraints, a paper-and-pencil administration was considered preferable to a computer based and/or online administration because it would limit the risk for skewed data resulting from participants’ disparate typing abilities. Another advantage of a paper-based instrument was that participants would be free to include diacritics, whereas they may be unfamiliar with how to do so on a keyboard. A concern, however, was that a written task in Spanish would be intimidating for HS who, unaccustomed to writing and reading in their HL,
tend to have lower levels of literacy and writing abilities when it comes to Spanish. Nevertheless, the benefits of using a paper-based instrument outweighed this concern. To help put the HS participants at ease, task instructions emphasized the inconsequential nature of spelling and accent marking, as well as the researcher’s interest in any and all words (formal, slang, vulgar, etc.) that came to mind. An initial pre-task period of informal mingling and refreshments also served to reduce any task-related apprehension.

The data collection sessions were initiated with a brief introduction of myself and the study, after which a set of oral preparatory instructions in Spanish were given:

You are going to receive a packet containing 22 sheets of paper. Each contains a list of instructions and a specific theme written at the top. The first sheet is for practice. The other 21 sheets form part of the official task. Once given the packet, please do not look ahead at any time unless directed.

I then distributed the packets to the participants. Eight packet versions, differing only in the order of the prompts, were distributed randomly. This was done to control for task effects in which the beginning prompts might be favored and the final prompts disfavored in terms of performance. Below the theme on each sheet were provided 40 possible lines to fill with lexical items. A sample task sheet is included in Appendix D. A practice round involved reviewing and clarifying specific task instructions, followed by a two-minute task based on a sample theme (emotions). For easy reference, the instructions (in Spanish) were repeated on each of the 22 sheets. They included the following:

*Write down all the words that come to mind related to the given prompt, from 1 and on. Please keep in mind the following:*

- You will have two minutes per topic.
- It is important to refrain from speaking with others during word production.
• Work with only one topic at a time. Do not look ahead. Wait for the sign from the investigator indicating it is OK to proceed to the next topic.
• You may write without regard for correct spelling. If you are unsure of a word’s spelling, please write the word anyway.
• If a word seems improper or vulgar, write it anyway.
• Do not write proper names.
• If you need extra space, continue your list on the blank section of the sheet.

After the two-minute time limit of the practice round expired, as a group we reviewed some acceptable answers. I purposefully offered some colloquial and vulgar terms (both adjectives and nouns), such as mañoso ‘cranky’ and cabreo ‘anger, rage’, among other more general or formal words. As well as providing a model for the array of acceptable word types, this potentially served to mitigate any unease felt by the participants.

Next, the 21 lexical categories were presented in succession. The categories were not announced aloud, as the randomized ordering meant each individual would be responding to a different topic at any given interval. The two-minute time allotments were marked by a buzzer, after which participants immediately turned the page and began with the next prompt. The lexical categories comprising the task are provided below in Table 3-4:
**TABLE 3-4**
**SEMANTIC DOMAINS AND CATEGORIES**

<table>
<thead>
<tr>
<th>Category</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home/Personal/Basic</td>
<td>(1) <em>El cuerpo humano</em> ‘The Human Body’</td>
</tr>
<tr>
<td></td>
<td>(2) <em>La ropa</em> ‘Clothing’</td>
</tr>
<tr>
<td></td>
<td>(3) <em>La casa</em> ‘Home’</td>
</tr>
<tr>
<td></td>
<td>(4) <em>Los muebles</em> ‘Furniture’</td>
</tr>
<tr>
<td></td>
<td>(5) <em>Comida y bebida</em> ‘Food and Drink’</td>
</tr>
<tr>
<td></td>
<td>(6) <em>Los colores</em> ‘Colors’</td>
</tr>
<tr>
<td></td>
<td>(7) <em>Espiritualidad y religión</em> ‘Spirituality and Religion’</td>
</tr>
<tr>
<td>Community</td>
<td>(8) <em>Medios de transporte</em> ‘Modes of Transportation’</td>
</tr>
<tr>
<td></td>
<td>(9) <em>Juegos y entretenimiento</em> ‘Games and Entertainment’</td>
</tr>
<tr>
<td></td>
<td>(10) <em>La ciudad</em> ‘The City’</td>
</tr>
<tr>
<td></td>
<td>(11) <em>El campo</em> ‘The Countryside’</td>
</tr>
<tr>
<td></td>
<td>(12) <em>La escuela</em> ‘School’</td>
</tr>
<tr>
<td></td>
<td>(13) <em>Profesiones y oficios</em> ‘Professions and Trades’</td>
</tr>
<tr>
<td></td>
<td>(14) <em>Los animales</em> ‘Animals’</td>
</tr>
<tr>
<td>Society</td>
<td>(15) <em>El medio ambiente</em> ‘The Environment’</td>
</tr>
<tr>
<td></td>
<td>(16) <em>La economía</em> ‘The Economy’</td>
</tr>
<tr>
<td></td>
<td>(17) <em>Las artes</em> ‘The Arts’</td>
</tr>
<tr>
<td></td>
<td>(18) <em>Ciencia y tecnología</em> ‘Science and Technology’</td>
</tr>
<tr>
<td></td>
<td>(19) <em>La política</em> ‘Politics’</td>
</tr>
<tr>
<td></td>
<td>(20) <em>Salud y enfermedades</em> ‘Health and Sicknessnes’</td>
</tr>
<tr>
<td></td>
<td>(21) <em>Problemas sociales</em> ‘Social Problems’</td>
</tr>
</tbody>
</table>

---

The participants' religious affiliations were not requested. However, it should be noted that because the exile-background participants were Marxists or secular liberals, they tend to be less religiously active (despite a Catholic upbringing and devout Catholic extended family members). Sweden's population is .01% Catholic (Catholic diocese of Stockholm 2013) while Chile's is 55% Catholic.
One of the principal aims of this study was to nuance the generalizing claims that HS lexicon is restricted to the often unspecified “home and family” domains. To study the range of domain-based lexical knowledge of HS in the community examined, a variety of domestic and extradomestic themes were selected to serve as prompts. Thirteen of the themes were drawn from the standard set of topics used in the Panhispanic Project of Available Lexicon (Samper Padilla & Samper Hernández 2007), which were originally adopted from the work of Gougenheim and his collaborators (1964) in France. Another six (Los colores ‘Colors’; Espiritualidad y religión ‘Spirituality and religion’; La política ‘Politics’; La economía ‘The economy’; Las artes ‘The arts’; El medioambiente ‘The environment’; Ciencia y Tecnología ‘Science and Technology’) were chosen to round out the domain continuum with more abstract, societal topics, as well as for the additional comparative value they could offer with lexical availability studies carried out in Chile (Valencia & Echeverría 1999; Valencia 2005). A final theme, Problemas Sociales ‘Social Problems’ was selected for its abstract and societal qualities.

For the analysis, in addition to interest in testing between-group differences on each of the 21 themes, the same were organized into three larger categories pursuant to their perceived degree of proximity to the home and family spheres. These supra-categories, each containing seven themes, appear in the leftmost column of Table 3-4. Under the assumption that the quantitative and qualitative attributes of HS vocabulary could vary as topics become less routine and more abstract, this condensed classification provided a framework through which to view possible landmark between-group differences along the continuum. The steps of the analytical process are discussed in the following sections.
3.4 Processing and Coding

As a first step to format the data for statistical and other analyses, the word lists underwent an editing process. This involved applying the protocol suggested by the Panhispanic Project as outlined by Samper Padilla et al. (2003). Alternate spellings across participants (e.g., *lavar*, *labar*), for example, or words offered by some participants in the singular and by others in the plural (e.g., *camisa*, *camisas*) were homogenized so that they would be recognized as tokens of the same type. However, these original-form productions were retained, as their variations are valuable for other research objectives. While it is understood that adhering to the norms of the RAE is fraught, the decision was made in order to permit comparisons with Project data from other dialect regions. Table 3-5 contains the editing guidelines followed.

**TABLE 3-5**
LEXICAL AVAILABILITY EDITING GUIDELINES

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Duplicated lexical items within a participant’s list for a particular topic are eliminated.</td>
</tr>
<tr>
<td>2.</td>
<td>As a general rule, spelling norms adhere to the RAE. In the case of borrowings, if they do not appear in the general dictionaries, the original spelling is used.</td>
</tr>
<tr>
<td>3.</td>
<td>Spelling homogenization in the cases of multiple accepted spellings, e.g., gineta / jineta. Decisions are based on the norms contained in the DRAE.</td>
</tr>
<tr>
<td>4.</td>
<td>Neutralization of flexive variants. The unmarked paradigm form is used: verbs in the infinitive; nouns and adjectives in masculine singular form. Exceptions: - game terms: veo, veo; policías y ladrones. - lexicalized diminutives: mesilla o mesita. - nouns with a verb or other noun complement: estante de libros - possible meaning differences: pecho - pechos.</td>
</tr>
<tr>
<td>5.</td>
<td>Full word forms are joined via parenthesis with their apocopate forms: boligrafo - boli &gt; boli(grafo); televisión - tele &gt; tele(visión).</td>
</tr>
<tr>
<td>6.</td>
<td>Full lexemes and their shortened variants are joined via parenthesis: columna - columna vertebral &gt; columna (vertebral).</td>
</tr>
<tr>
<td>7.</td>
<td>Trade brands: As a general rule, the only ones retained are those that have undergone full lexicalization: coca-cola. Caution must be used, given the difficulty of determining an item’s degree of lexicalization.</td>
</tr>
</tbody>
</table>

As part of the editing phase, all words were entered into a spreadsheet organized by participant, and the words per participant quantified. These figures were recorded alongside the codified
questionnaire and proficiency measure data. These data were then subjected to statistical analyses using R and SPSS 18. To address RQ1a, Tukey HSD pairwise comparisons were performed in R through a linear mixed model with the number of words produced as the dependent variable and speaker generation (n=2), semantic theme (n=21), and category type (n=3) as independent variables. In addition, speaker was included as a random effect. The Tukey HSD was chosen for its ability to account for the unequal sample sizes between the two groups and because it includes a correction for multiple comparisons at the level of the test statistic. To address RQ1b, Tukey HSD pairwise comparisons were again realized in R with number of words produced as the dependent variable and gender, birth order, parental education, and exile background as independent variables; the effect of these variables on word production in each domain and supra-category was considered. To address RQ1c, bivariate correlations were performed in SPSS to test the relationship between the speakers’ total word production (across all domains) and their data for the following variables: HL instruction, Spanish language use, Spanish proficiency, English proficiency, Spanish proficiency self-rating, English proficiency self-rating, Spanish-language media exposure, English-language media exposure.

In order to address research questions 2a-d, the word lists for six semantic domains were entered into a database (www.dispolex.com) and their frequency and availability indexes (AI) computed. Two domains were selected randomly from each supra-category in order to form a representative sample of the domain continuum. These included The Human Body, Furniture, Modes of Transportation, Professions and Trades, The Economy, and Politics. The online program relies on a formula developed by López Chávez and Strassburger (1991) which calculates lexical availability combining three principal data points: (1) the total number of participants, (2) the number of participants that produced a given word (i.e., its frequency), and
Accordingly, a lexical item’s AI is high if it is generated by many speakers early on in their lists. Conversely, a lexical item has a low AI if it is produced by few speakers late in their lists. Also, given that the primary objective of lexical availability research is to determine the most available lexicon (understood as that which is offered most readily), the formula is designed to assign increasingly less importance to frequency going from the first word on the list to the last (López Morales 1986; Etxebarria 2014). The formula is:

\[ D(P) = \sum_{i=1}^{n} e^{-2.3 \times \left( \frac{i-1}{n-1} \right)} \times \frac{f}{I_1} \]

Hernández Múnoz et al. (2006, p. 737) describe it in the following terms: "\(D(P)\) = the lexical availability of the word \(j\) for a given category; \(I\) = the total number of participants who completed the task; \(i\) = the position of word \(j\) in a given list; \(f\) = the number of participants who wrote word \(j\) at that position in their list; \(n\) = the lowest position occupied by word \(j\) in any list produced for the category in question; and \(e\) = the natural number (2.718181818459045)." The program generates a list-form output organized by domain in which all words appear in descending order according to their AI. Alongside a word’s AI is also displayed its simple group frequency. Portions of these availability lists can then be used to examine different questions. Figure 3-1 provides an example of an output list (only the first 15 available words are shown).

Figure 3-1. Sample lexical availability output list
To address RQ2a, the program was used to perform between-group lexical compatibility analyses on a subset of the groups’ productions in each domain. Specifically, the types considered were those with an availability index of .02 or higher (representing roughly 80% of all domain tokens). This action excluded the bulk of late-mentioned, idiosyncratic lexical items, permitting a focus on the words most representative of the groups’ lexicons (Samper Padilla & Samper Hernández 2007). The compatibility analyses were realized as a first step that would provide a general quantitative observation of rate of G1-G2 lexical correspondence, i.e. the percent overlap between the groups' lexical productions in each domain. As a continuation, it was of interest to determine which words constituted those that differed between the groups, from which generational patterns could be discerned. For this, I used an online tool called Venny (bioinfogp.cnb.csic.es/tools/venny) to perform a Venn-diagram-like analysis of the Top 50 words of the lists corresponding to each of the six semantic domains. This enabled an analysis of the types of lexical items corresponding to three word groups: shared; unique to G1; unique to G2.
A key factor in lexical compatibility across datasets is lexical diversity. To obtain a closer look at the lexical diversity of each group, the Guiraud’s Index (Root TTR) was computed for each domain. That is, the measure was used to calculate the level of homogeneity in each group’s domain responses. This permitted comparisons of how groups differed by theme in terms of lexical diversity. Root TTR was chosen because of its greater reliability when comparing texts of different lengths (in this case, of different discrete word totals).

To address RQ2b-c, the full output lists for the six domains were used so that both integrated (produced by more than one participant) and spontaneous (produced by a single participant) borrowings would be accounted for. Once tallied, the number of Swedish and English borrowings were each divided by the total number of words of the data subset (the six domains) from which they were drawn. This resulted in a “density index” (Alba 1995b; Hernández Muñoz 2009) for each speaker group representing the overall concentration of each class of borrowings in the data. The same was performed at the level of the domain, permitting a look at the distributional patterns of contact-induced lexicon with regard to semantic context. To complement the density indexes obtained, I also divided the overall token counts of each of the borrowing types (Swedish and English) by the total number of speakers, which generated a mean number of borrowings per speaker. This figure is intended to represent the vitality of the contact lexicon in the speech community as a whole and across generations (rather than their density in the datasets) (Hernández Muñoz 2009). The Swedish and English contact lexicon was subsequently classified and tallied according to borrowing type (see Table 2-4). These totals were then examined comparatively to determine whether notable patterns existed cross-generationally and between the two sources of borrowings.
With respect to the presence of Peninsular Spanish lexical items, a type/token analysis was performed cross-generationally across all six domains. To accurately determine which items were uniquely from Spain and not found in Chile, I relied on the *Diccionario de Americanismos* (Asociación de Academias de la Lengua Española), Valencia and Echeverría's (1999) homeland lexical availability data, as well as the *Diccionario de la Real Academia Española* (DRAE).

To address RQ2d, subsets of the AI lists for the same six domains were again used. For this analysis, the Top 15 lexical items (according to AI) were considered. Again, a reduction in the amount of words pondered is helpful because it excludes idiosyncratic responses (offered by a single individual), which tend to be less representative of the group’s lexical store. RQ2d was concerned with the category conceptualizations that might be discernable in the rankings of the most available items of a given domain, as well as in the absence versus presence of words in these top positions. Side-by-side between-group comparisons were performed under the assumption that an AI reflects the prototypical value of a lexical item within a semantic domain; the higher the AI, the more prototypical the lexical item is of that domain. These analyses considered not only disparate prototypicalities of individual words across generations, but also the presence in the lists of differing collections of semantically related words that, together, were revelatory of a group’s domain conceptualizations.

Table 3-6 provides a brief summary of the RQs and their accompanying analyses:
<table>
<thead>
<tr>
<th>RQ</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a. Quantitative differences between G1 and G2 cross-domain productions.</td>
</tr>
<tr>
<td></td>
<td>b. Relationship between G2 lexical production and social variables.</td>
</tr>
<tr>
<td></td>
<td>c. Relationship between lexical production and language variables.</td>
</tr>
<tr>
<td>2</td>
<td>a. Between-group lexical compatibility rates and characteristics of divergent lexicon.</td>
</tr>
<tr>
<td></td>
<td>b. Between-group density and types of Swedish and English borrowings</td>
</tr>
<tr>
<td></td>
<td>c. Between-group density and types of Peninsularisms</td>
</tr>
<tr>
<td></td>
<td>d. Between-group thematic conceptualizations.</td>
</tr>
</tbody>
</table>
4. RESULTS AND DISCUSSION

4.1 Research Question 1a: Quantitative Comparisons

A total of 15,577 tokens were produced by the 53 participants across the 21 semantic domains, with 8,817 corresponding to the G1 (n=30) and 6,760 to the G2 (n=23). This represents an identical per speaker average of 294 total words for both groups. The average mean per domain was approximately 14 words for both groups, surpassing the 10-item average documented for non-native learners of Spanish (Samper Hernández 2001). Table 4-1 displays each group’s ranking of the semantic domains according to the mean number of words produced.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Domain</th>
<th>G1 (n=30)</th>
<th>Mean</th>
<th>Domain</th>
<th>G2 (n=23)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Human Body</td>
<td></td>
<td>18.50</td>
<td>The Human Body</td>
<td></td>
<td>20.13</td>
</tr>
<tr>
<td>2</td>
<td>Countryside</td>
<td></td>
<td>17.43</td>
<td>Animals</td>
<td></td>
<td>18.48</td>
</tr>
<tr>
<td>3</td>
<td>Animals</td>
<td></td>
<td>17.17</td>
<td>School</td>
<td></td>
<td>17.78</td>
</tr>
<tr>
<td>4</td>
<td>School</td>
<td></td>
<td>16.80</td>
<td>Food and Drink</td>
<td></td>
<td>17.17</td>
</tr>
<tr>
<td>5</td>
<td>Food and Drink</td>
<td></td>
<td>16.50</td>
<td>Spirituality and Religion</td>
<td></td>
<td>16.65</td>
</tr>
<tr>
<td>6</td>
<td>Home</td>
<td></td>
<td>16.37</td>
<td>Clothing</td>
<td></td>
<td>16.52</td>
</tr>
<tr>
<td>7</td>
<td>Clothing</td>
<td></td>
<td>16.23</td>
<td>Home</td>
<td></td>
<td>15.74</td>
</tr>
<tr>
<td>8</td>
<td>Professions and Trades</td>
<td></td>
<td>15.63</td>
<td>City</td>
<td></td>
<td>14.57</td>
</tr>
<tr>
<td>9</td>
<td>City</td>
<td></td>
<td>15.47</td>
<td>Colors</td>
<td></td>
<td>14.87</td>
</tr>
<tr>
<td>10</td>
<td>Colors</td>
<td></td>
<td>14.83</td>
<td>Health and Sickness</td>
<td></td>
<td>13.96</td>
</tr>
<tr>
<td>11</td>
<td>Spirituality and Religion</td>
<td></td>
<td>13.80</td>
<td>Countryside</td>
<td></td>
<td>13.78</td>
</tr>
<tr>
<td>12</td>
<td>Health and Sickness</td>
<td></td>
<td>12.80</td>
<td>Games and Entertainment</td>
<td></td>
<td>13.39</td>
</tr>
<tr>
<td>13</td>
<td>Modes of Transportation</td>
<td></td>
<td>12.77</td>
<td>Modes of Transportation</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>Games and Entertainment</td>
<td></td>
<td>12.70</td>
<td>Professions and Trades</td>
<td></td>
<td>12.48</td>
</tr>
<tr>
<td>15</td>
<td>Furniture</td>
<td></td>
<td>12.67</td>
<td>The Economy</td>
<td></td>
<td>12.13</td>
</tr>
<tr>
<td>16</td>
<td>The Economy</td>
<td></td>
<td>12.14</td>
<td>Politics</td>
<td></td>
<td>11.65</td>
</tr>
<tr>
<td>17</td>
<td>The Environment</td>
<td></td>
<td>11.77</td>
<td>Furniture</td>
<td></td>
<td>11.43</td>
</tr>
<tr>
<td>18</td>
<td>Politics</td>
<td></td>
<td>11.60</td>
<td>Science and Technology</td>
<td></td>
<td>10.65</td>
</tr>
<tr>
<td>19</td>
<td>Social Problems</td>
<td></td>
<td>10.77</td>
<td>The Environment</td>
<td></td>
<td>10.57</td>
</tr>
<tr>
<td>20</td>
<td>The Arts</td>
<td></td>
<td>9.83</td>
<td>The Arts</td>
<td></td>
<td>10.22</td>
</tr>
<tr>
<td>21</td>
<td>Science and Technology</td>
<td></td>
<td>9.63</td>
<td>Social Problems</td>
<td></td>
<td>9.17</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>14.04</td>
<td></td>
<td></td>
<td>14.02</td>
</tr>
</tbody>
</table>

Table 4-1: Domain ranking by speaker group according to mean number of words produced.
For the G1, the domains *The Human Body* and *Countryside* garnered the most responses, while the fewest were generated in the areas of *The Arts* and *Science and Technology*. Meanwhile, the G2 produced on average the most words in the domains of *The Human Body* as well, and in *Animals*, while they produced the fewest under *Social Problems* and, like the G1, under *The Arts*.

These rankings clearly show that, due to their variable semantic nature, some themes lend themselves to greater response generation than others. As discussed in Chapter 2, research supports the idea that regardless of language background (L2, monolingual, heritage), there is a high probability that speakers will generate more vocabulary for a basic domain like *The Human Body* than for a more complex domain such as *The Economy*. Although this pattern was generally confirmed cross-generationally, a few important ranking differences were observed, which are shaded in Table 4-1 for ease of identification. The largest concerned *Countryside*, for which the G1 had a considerably higher ranking (#2) than the G2 (#11). The domain *Professions and Trades* also ranked notably higher for the G1 (#8, vs. #14 for the G2), and *Spirituality and Religion* for the G2 (#5, vs. #11 for the G1). The means and mean differences are visually presented in Figures 4-1 and 4-2.
Figure 4-1. Group means along the domain continuum
Figure 4-2.  G1 vs. G2 mean differences
Table 4-2 provides further pairwise statistical data for all 21 domains:

### TABLE 4-2
PAIRWISE COMPARISONS OF GROUP MEANS

<table>
<thead>
<tr>
<th>Domain</th>
<th>G1 (n=30)</th>
<th>G2 (n=23)</th>
<th>Difference</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Body</td>
<td>18.50 5.98</td>
<td>20.13 7.58</td>
<td>-1.63</td>
<td>-0.99</td>
<td>0.3245</td>
</tr>
<tr>
<td>Clothing</td>
<td>16.23 5.30</td>
<td>16.52 6.69</td>
<td>-0.29</td>
<td>-0.17</td>
<td>0.8616</td>
</tr>
<tr>
<td>Home</td>
<td>16.37 6.08</td>
<td>15.74 7.23</td>
<td>0.63</td>
<td>0.38</td>
<td>0.7045</td>
</tr>
<tr>
<td>Furniture</td>
<td>12.67 3.79</td>
<td>11.43 4.19</td>
<td>1.24</td>
<td>0.74</td>
<td>0.4566</td>
</tr>
<tr>
<td>Food and Drink</td>
<td>16.50 6.23</td>
<td>17.17 7.87</td>
<td>-0.67</td>
<td>-0.73</td>
<td>0.4656</td>
</tr>
<tr>
<td>Colors</td>
<td>14.83 5.78</td>
<td>14.87 4.58</td>
<td>-0.04</td>
<td>-0.02</td>
<td>0.9825</td>
</tr>
<tr>
<td>Spirituality and Religion</td>
<td>13.80 6.10</td>
<td>16.65 6.44</td>
<td>-2.85</td>
<td>-1.72</td>
<td>0.0849*</td>
</tr>
<tr>
<td>Modes of Transportation</td>
<td>12.77 5.10</td>
<td>13.00 5.25</td>
<td>-0.23</td>
<td>-0.14</td>
<td>0.8878</td>
</tr>
<tr>
<td>Games and Entertainment</td>
<td>12.70 5.36</td>
<td>13.39 6.24</td>
<td>-0.69</td>
<td>-0.42</td>
<td>0.6761</td>
</tr>
<tr>
<td>City</td>
<td>15.47 6.59</td>
<td>14.57 7.69</td>
<td>0.9</td>
<td>0.54</td>
<td>0.5859</td>
</tr>
<tr>
<td>Countryside</td>
<td>17.43 7.50</td>
<td>13.78 5.92</td>
<td>3.65</td>
<td>2.21</td>
<td>0.0275**</td>
</tr>
<tr>
<td>School</td>
<td>16.80 6.62</td>
<td>17.78 7.14</td>
<td>-0.98</td>
<td>-0.59</td>
<td>0.5526</td>
</tr>
<tr>
<td>Professions and Trades</td>
<td>15.63 5.64</td>
<td>12.48 5.4</td>
<td>3.15</td>
<td>1.91</td>
<td>0.0567*</td>
</tr>
<tr>
<td>Animals</td>
<td>17.17 4.31</td>
<td>18.48 6.33</td>
<td>-1.31</td>
<td>-1.14</td>
<td>0.2564</td>
</tr>
<tr>
<td>Environment</td>
<td>11.77 5.91</td>
<td>10.57 5.83</td>
<td>1.2</td>
<td>0.73</td>
<td>0.4678</td>
</tr>
<tr>
<td>Economy</td>
<td>12.14 5.89</td>
<td>12.13 5.85</td>
<td>0.01</td>
<td>-0.11</td>
<td>0.9132</td>
</tr>
<tr>
<td>Arts</td>
<td>9.83 3.91</td>
<td>10.22 5.20</td>
<td>-0.39</td>
<td>-0.23</td>
<td>0.8164</td>
</tr>
<tr>
<td>Science and Technology</td>
<td>9.63 3.96</td>
<td>10.65 5.61</td>
<td>-1.02</td>
<td>-0.35</td>
<td>0.7241</td>
</tr>
<tr>
<td>Politics</td>
<td>11.60 4.69</td>
<td>11.65 6.22</td>
<td>-0.05</td>
<td>-0.03</td>
<td>0.9748</td>
</tr>
<tr>
<td>Health and Sickness</td>
<td>12.80 5.44</td>
<td>13.96 6.31</td>
<td>-1.16</td>
<td>-0.7</td>
<td>0.4846</td>
</tr>
<tr>
<td>Social Problems</td>
<td>10.77 5.39</td>
<td>9.17 5.42</td>
<td>1.6</td>
<td>0.96</td>
<td>0.3358</td>
</tr>
<tr>
<td>Overall</td>
<td>14.04 5.50</td>
<td>14.02 6.04</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: For those domains in which group means differed significantly, the row is shaded for ease of identification.
*Approached significance at the .05 level; **Significant at the .05 level.

Although the groups’ overall mean number of words per category was equal (14.04 versus 14.02), the higher G2 standard deviation (SD) (6.04, versus 5.5 for the G1) indicated somewhat greater variation among this group. The heterogeneity of HS abilities, given their diverse life histories and experience with the language, is a reality well documented in the literature (cf. Kondo-Brown 2004, 2005; Mikulski 2010; Parada 2013), as discussed in Chapters 1 and 2. For both groups the highest means were found toward the beginning of the basic–abstract domain spectrum and the lowest means toward the end. In other words, for both groups, as topic...
sophistication increased, word production tended to decrease. Because the prompts were administered in a randomized fashion, this pattern could not be interpreted as a task effect.

Given the G2 speakers’ comparatively limited exposure to and use of Spanish in educational and professional contexts of society, it was predicted that their lexical productivity would diverge more drastically from that of the G1 in the extradomestic domains regarded as more conceptually complex. However, a cursory look at the data reveals highly comparable group means in most domains. The results of a linear mixed model with number of words produced as the dependent variable, with speaker generation (n=2), semantic domain (n=21), domain supra-category (n=3) as fixed effects, and with speaker as a random effect, revealed a main effect of semantic domain \( F(2, 1055)=25.62, p<.01 \) and of domain supra-category \( F(1, 1019)=126.25, p<.01 \), as well as an interaction between speaker generation and semantic domain \( F(20, 1019)=2.15, p<.05 \). No interaction was found between speaker generation and domain supra-category. Tukey pairwise comparisons indicated that only for the domain *The Countryside* did the G1 produce significantly more words (an average of 3.65 more) than the G2 \( [p<.05] \). The mean differences for *Professions and Trades* (3.15) and *Spirituality and Religion* (-2.85) approached significance, with the last indicating a G2 advantage \( [p=.057; p=.085] \). The prediction that generational divergences would be greatest among the most sophisticated topics thus does not obtain. These results also show that although there are certain significant domain-level differences in word production rates, the production patterns from the most domestic to least domestic areas of the lexicon are generally highly similar between the G1 and the G2.

It is interesting that the two domains for which G2 speakers produced substantially fewer words (*The Countryside* and *Professions and Trades*) are located in the center of the basic-abstract (or domestic-extradomestic) continuum, and appear to represent rather mundane
vocabulary with which HS might be expected to be familiar. One explanation for these particular disparities may be that because of the somewhat closed and metonymic nature of these categories, as opposed to the more open and abstract *The Economy* and other topics at the far end of the spectrum, there is a reduced ability to resort to peripheral vocabulary in the face of lexical gaps. One might also question whether there is something about these specific domains in the groups’ distinct upbringings that could explain the divergent results. For example, in the case of *The Countryside*, it may be that the urban upbringing of the G2 in Sweden has provided few opportunities for the development of related lexicon, whereas the G1 speakers’ more likely exposure to such lexical content through past residential, educational, and professional experiences in monolingual settings has resulted in a larger inventory. Likewise, it is possible that the G2’s greater exposure to a range of spiritual and religious ideologies and practices in multicultural Sweden explains their higher productivity under the topic of *Spirituality and Religion*. The groups’ actual productions will be examined later in the chapter, with the aim of further elucidating the between-group differences in these and other domains.

A look at the participants’ productions on different collections of domains (i.e., supra-categories) provides a clearer picture of the groups’ shared linear pattern: overall, the most numerous responses pertain to Basic; Community has fewer; and there is a substantial drop in Society (Table 4-3). The main effect found in the statistical model associated with this pattern lends credibility to the study's organization of the domains according to personal/domestic relevance and level of sophistication. We can recall from the introduction and review of the literature that previous descriptions of HS lexicon as bound to the home and family contexts

---

17 Although this information was not quantified, it could be gleaned from the questionnaire data that many G1 had lived or worked in rural areas.
failed to adequately outline which domains should be classified as such; these results offer empirical data to assess and nuance those descriptions.

Unlike the domain comparisons, the model did not find significant differences between the group means in the three supra-categories. Table 4-3 shows that the means did differ somewhat, however, with the largest difference favoring the G2 in Basic (4.15 more words produced on average). The G2's higher means in five of the seven Basic domains explains this difference, and points to their extremely robust lexical knowledge in the domestic realm.

TABLE 4-3
BETWEEN-GROUP SUPRA-CATEGORY MEAN COMPARISONS

<table>
<thead>
<tr>
<th>Supra-category</th>
<th>G1</th>
<th>G2</th>
<th>Difference, Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>SD/ Mean %</td>
</tr>
<tr>
<td>Basic</td>
<td>108.90</td>
<td>101.5</td>
<td>44</td>
</tr>
<tr>
<td>Community</td>
<td>107.97</td>
<td>100.5</td>
<td>25</td>
</tr>
<tr>
<td>Society</td>
<td>78.13</td>
<td>74.5</td>
<td>50</td>
</tr>
<tr>
<td>Range</td>
<td>30.77</td>
<td>27</td>
<td>34.17</td>
</tr>
</tbody>
</table>

As for Community and Society, the G1 had a higher word average, though negligible in the latter case (0.22 more words). The higher G1 mean (4.49 more words) for Community was primarily due to the speakers’ greater productivity in the categories of *The Countryside* and *Professions and Trades* for which, as discussed previously, the G1 had significantly higher means. A look at the medians (also displayed in Table 4-3), an arguably more telling figure, reveals slightly more marked (a seven word advantage for the G2 in Basic) and even reversed (G2 advantage in Community, G1 advantage in Society) between-group differences. However, because these differences did not achieve statistical significance, we can again conclude that domain complexity does not predict inter-generational difference in quantitative aspects of lexical
production in the context of a lexical availability task. Taken together, the quantitative results confirm that the G2 participants know an impressive number of words pertaining to a variety of semantic domains, and that they are able to recall and produce them with great ease.

A final observation with respect to the quantitative data in Table 4-3 has to do with within-group performance variability. The range of change in means and medians for each group from Basic to Society revealed a larger spectrum for the G2 (34.17 and 39 word ranges, respectively) than for the G1 (30.77 and 27 word ranges). This again seems to point to the particular heterogeneity of the G2 with respect to HL proficiency. However, more nuanced patterns of response heterogeneity emerge when considering the SD in relation to the mean for each supra-category. This relationship is represented in Table 4-3 as a percentage for ease of comparison between the groups.

We can observe that in Basic, the G1 participants’ word totals were actually more varied (the SD represented 44% of the mean) than those of the G2 (the SD represented 39% of the mean). Alternatively, for Community, the G2 showed exceedingly more variability across speakers with the SD representing 43% of the mean, versus 25% for the G1. Lastly, for Society, participant variability between the groups was similar (52% versus 50%). Thus, Community stands out as the supra-category in which the variability of lexical knowledge across G2 participants diverges most from that of the G1. Such inconsistency elicits questions as to what about the nature of this supra-category and the domains it comprises is responsible for the G2’s much higher performance variability. That is, what makes the lexicon of these domains so universally familiar to the G1 yet so variably familiar to the G2? A look at the individual SDs of the domains comprising Community revealed that *Animals, City,* and *Professions and Trades* were the thematic areas of greatest difference in terms of G1 vs. G2 within-group variability.
While extensive vocabulary in these domains appears to be more of a given for the G1, a G2's lexical breadth in these areas seems highly dependent on input conditions.

Table 4-4 presents the mean comparisons on 14 domains between the two participant groups and monolingual homeland youth (aged 18, approximately) from Valencia & Echeverria (1999). The 14 domains included those that were tested in both studies.

### Table 4-4

**COMPARISON OF MEANS WITH HOMELAND SPEAKERS**

<table>
<thead>
<tr>
<th>Domain</th>
<th>G1 Chilean Swedes</th>
<th>G2 Chilean Swedes</th>
<th>Homeland Youth&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Difference G1-Homeland</th>
<th>Difference G2-homeland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Body</td>
<td>18.50</td>
<td>20.13</td>
<td>21.39</td>
<td>-2.89</td>
<td>-1.26</td>
</tr>
<tr>
<td>Clothing</td>
<td>16.23</td>
<td>16.52</td>
<td>20.30</td>
<td>-4.07</td>
<td>-3.78</td>
</tr>
<tr>
<td>Home</td>
<td>16.37</td>
<td>15.74</td>
<td>22.86</td>
<td>-6.49</td>
<td>-7.12</td>
</tr>
<tr>
<td>Furniture</td>
<td>12.67</td>
<td>11.43</td>
<td>11.06</td>
<td>1.61</td>
<td>0.37</td>
</tr>
<tr>
<td>Food and Drink</td>
<td>15.97</td>
<td>17.17</td>
<td>23.84</td>
<td>-7.87</td>
<td>-6.67</td>
</tr>
<tr>
<td>Spirituality &amp; Religion</td>
<td>13.80</td>
<td>16.65</td>
<td>18.15</td>
<td>-4.35</td>
<td>-1.5</td>
</tr>
<tr>
<td>Modes of Transportation</td>
<td>12.77</td>
<td>13</td>
<td>16.84</td>
<td>-4.07</td>
<td>-3.84</td>
</tr>
<tr>
<td>Games &amp; Entertainment</td>
<td>12.70</td>
<td>13.39</td>
<td>16.02</td>
<td>-3.32</td>
<td>-2.63</td>
</tr>
<tr>
<td>Professions &amp; Trades</td>
<td>15.63</td>
<td>12.48</td>
<td>18.97</td>
<td>-3.34</td>
<td>-6.49</td>
</tr>
<tr>
<td>Environment</td>
<td>11.77</td>
<td>10.57</td>
<td>11.14</td>
<td>0.63</td>
<td>-0.57</td>
</tr>
<tr>
<td>Economy</td>
<td>12.14</td>
<td>12.13</td>
<td>14.92</td>
<td>-2.78</td>
<td>-2.79</td>
</tr>
<tr>
<td>Arts</td>
<td>9.83</td>
<td>10.22</td>
<td>14.80</td>
<td>-4.97</td>
<td>-4.58</td>
</tr>
<tr>
<td>Science &amp; Technology</td>
<td>9.63</td>
<td>10.22</td>
<td>12.42</td>
<td>-2.79</td>
<td>-2.2</td>
</tr>
<tr>
<td>Politics</td>
<td>11.60</td>
<td>11.65</td>
<td>15.02</td>
<td>-3.42</td>
<td>-3.37</td>
</tr>
<tr>
<td>Overall</td>
<td>13.54</td>
<td>13.66</td>
<td>17</td>
<td>-3.46</td>
<td>-3.34</td>
</tr>
</tbody>
</table>

Note: Shaded areas are those G1 or G2 means that exceeded the corresponding Homeland means.

<sup>a</sup> Valencia & Echeverria (1999), n=2052

While the G1 and G2 overall means were comparable (13.54 and 13.66 respectively), the homeland youth speakers produced an average of 17 words per domain. Thus, the G1 and G2 production rates more closely resemble each other than they do the rates of homeland speakers, who, based on their age, still had largely unspecialized vocabularies. The G1 – homeland difference is noteworthy, given that they share a monolingual upbringing. This result should be taken with caution given the unequal sample sizes, but may suggest attrition on the part of the
G1. Alternatively, age effects may be a factor; as discussed in Chapter 2, we can recall that Urrutia’s (2001) young adult participant group produced more lexicon than his middle aged and elderly groups, though differences were subtler than in the present data. Among the 14 domains compared, only for The Enviroment and Furniture did the G1 have a higher word average than the homeland speakers. Curiously, the G2 as well produced a slightly higher word average only in Furniture. An examination of the groups’ actual productions later in this chapter may shed light on these discrepancies; one might wonder, for example, whether the urban existence of Chilean Swedes plays a role in these particular domain advantages.

Lastly, it was also of interest to compare the participant data with results reported by Moreno-Fernández (2007) for U.S. Latino youth in Chicago. Since the author only provided the number of types (not tokens) his participants produced as a group in each domain, I have derived a per-speaker mean from those figures. After computing the same for the present study data (mean number of types per speaker) in four coinciding domains (Human Body, Furniture, Modes of Transportation, Professions and Trades)\(^\text{18}\), I performed side-by-side comparisons. Because the unit of analysis in this comparison was not tokens but types, it actually gave a better indication of lexical diversity than vocabulary size. That is, these data show how many unique words each participant on average contributed to the overall type total, and, thus, how diverse the responses were within each speaker group. The results are presented in Table 4-5:

---
\(^{18}\) As a subsample, these four domains were selected randomly from among the 13 overlapping ones between the two studies.
TABLE 4-5
COMPARISON OF MEANS WITH U.S. HERITAGE SPEAKERS

<table>
<thead>
<tr>
<th>Domain</th>
<th>G1 Chilean Swedes n = 30</th>
<th>G2 Chilean Swedes n = 23</th>
<th>U.S. Latino Youtha n = 48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Body</td>
<td>2.33</td>
<td>6.83</td>
<td>2.9</td>
</tr>
<tr>
<td>Furniture</td>
<td>1.87</td>
<td>2.04</td>
<td>2.5</td>
</tr>
<tr>
<td>Modes of Transportation</td>
<td>1.7</td>
<td>2.17</td>
<td>2.02</td>
</tr>
<tr>
<td>Professions and Trades</td>
<td>2.17</td>
<td>2.65</td>
<td>4.54</td>
</tr>
<tr>
<td>Overall</td>
<td>2.02</td>
<td>3.42</td>
<td>2.99</td>
</tr>
</tbody>
</table>

The data paint an interesting picture: both HS groups have a notably larger types-per-speaker average than the G1, meaning that their responses were more varied as a group than those of the G1.

Although generally in analyses of individual learner/speaker texts, higher measures of lexical diversity typically indicate stronger command of the target language or register, in the collective word lists of lexical availability (association) tasks, which are used to examine the lexicon of a speaker group, highly diverse responses across group members may indicate the exact opposite. This is because highly variable words produced in response to a thematic prompt can signal poorly organized, underdeveloped mental lexicons (Schmitt 2010). In a lexical availability task, response variability is constrained by semantic relevance to the given prompt and also register (effects of a written task), two areas in which HS may have weaker abilities, possibly explaining their less uniform word production across speakers. These data (in Table 4-5) suggest that HS have sizeable vocabularies, but a somewhat limited consensus among themselves (relative to the G1) with regard to the associations produced by a thematic prompt. The notion of lexical diversity will be further discussed in Section 4.3, which includes results of the qualitative analyses of the participants’ productions.
4.2 Research Questions 1b and 1c: Social and Language Factors

Pairwise comparisons were performed to test the relationship of four social factors with the G2 productions on the 21 semantic domains as well as the three supra-categories. For each factor, statistically significant differences were found in at least one domain. Table 4-6 displays the results.

<table>
<thead>
<tr>
<th>Social Variable</th>
<th>n</th>
<th>Total Words Mean</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td>345.13</td>
<td></td>
</tr>
<tr>
<td><strong>Birth Order</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oldest</td>
<td>9</td>
<td>272.89</td>
<td>Middle &gt; Oldest Human Body**, City*, School*, Politics*, Community*</td>
</tr>
<tr>
<td>Middle</td>
<td>8</td>
<td>309.5</td>
<td>Middle &gt; Youngest Human Body*</td>
</tr>
<tr>
<td>Youngest</td>
<td>6</td>
<td>306.3</td>
<td></td>
</tr>
<tr>
<td><strong>Parental Education Average</strong></td>
<td>3.73</td>
<td>281.54</td>
<td>University + &gt; University – Food&amp;Drink</td>
</tr>
<tr>
<td>University -</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University +</td>
<td>10</td>
<td>311</td>
<td></td>
</tr>
<tr>
<td><strong>Exile Background</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>15</td>
<td>331.6</td>
<td>Exile Yes &gt; Exile No Human Body, Food&amp;Drink, Colors, Spirituality&amp;Religion, Games&amp;Entertainment, Countryside, School, Professions&amp;Trades, Animals, Arts, Science&amp;Tech, Health&amp;Sickness, Basic, Community, Society</td>
</tr>
<tr>
<td>No</td>
<td>8</td>
<td>224.5</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05; **p < .01

With respect to gender, the male participants produced significantly more words in six domains from across the domestic-extradomestic continuum: Human Body, Home, Spirituality and
Religion, City, Professions and Trades, Science and Technology, and Politics. The mean difference was especially marked for Politics (significant at the .01 level). With respect to the supra-categories, male participants produced significantly more words in Community and Society. These results confirm the prediction, based on previous research in the community showing greater ethnic heritage affiliation on the part of males (Ganuza and King 2005), that male speakers would exhibit a proficiency advantage over their female counterparts.

For the factor of birth order, the prediction was that speakers of higher birth order would generally produce more words than those of lower birth order, especially from the third-born onward in light of research suggesting that the oldest and second-born often look very similar in terms of linguistic proficiency (Parada 2013). The results provide very little support for this prediction in that the middle children only produced significantly more than the youngest children in a single domain (Human Body). Also, surprisingly, the middle children produced significantly more lexicon than the oldest children in four domains and in the supra-category Community.

In terms of parental education, with an average of 3.73 out of 5 (1=primary education, 5=advanced degree), the G2 participants’ parents were fairly well educated. For the pairwise comparison, the participants were divided into two groups based on whether their parental education average was below or above 4 (university level). Although the total word means differed substantially in favor of speakers whose parents had obtained more formal education (311 versus 281.54 words), only for the domain Food and Drink was the mean difference returned as significant. Thus, the prediction that parental education would be significantly correlated with lexical knowledge was not substantiated by the data.
Finally, the factor of exile background generated several significant results, all in the predicted direction of an advantage for HS whose parents arrived in Sweden under political asylum. In all three supra-categories speakers of an exile background produced a significantly higher average number of words, as well as in 12 domains: *Human Body, Food and Drink, Colors, Spirituality and Religion, Games and Entertainment, Countryside, School, Professions and Trades, Animals, Arts, Science and Technology, and Health and Sicknesses*. It is important to note, however, that there was no significant numerical difference for *Politics*.

I now move on to the language factors considered in the analysis, Tables 4-7 and 4-8 summarize the applicable group means.

<table>
<thead>
<tr>
<th>TABLE 4-7</th>
<th>SUMMARY MEANS OF LANGUAGE VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G1 Mean (SD)</td>
</tr>
<tr>
<td>HL Education (school years)</td>
<td></td>
</tr>
<tr>
<td>Spanish proficiency test (% right)</td>
<td>95 (.06)</td>
</tr>
<tr>
<td>English proficiency test (% right)</td>
<td></td>
</tr>
<tr>
<td>Spanish use (20 questions, 5 pt. scale)</td>
<td>3.72 (0.73)</td>
</tr>
<tr>
<td>Spanish media exposure (3 questions, 5 pt. scale)</td>
<td>3.43 (1.28)</td>
</tr>
<tr>
<td>English media exposure (3 questions, 5 pt. scale)</td>
<td>2.64 (1.11)</td>
</tr>
</tbody>
</table>

As shown in Table 4-7, most of the G2 individuals reported a history of some formal study of the HL during their elementary and middle school years, though the amount of seat time varied widely (from 0 to 9 years). The results of the DELE Spanish proficiency measure showed
significantly greater accuracy on the part of the G1, though an 83% mean for the G2 also evidenced the heritage speakers’ robust knowledge.

With respect to the English proficiency measure, the participants knew an average of 57% of the words tested; however, there was an unexpectedly large degree of variation in the group (SD = 21.7) surpassing that of the corresponding Spanish proficiency measure. The results of the Spanish use questionnaire indicated that the G1 speakers use Spanish significantly more than the G2 in an array of daily activities and personal encounters. This pattern was mirrored in the group responses regarding regular media exposure to Spanish (TV, music, reading), though the smaller between-group difference here may be attributed to the G2 speakers' apparent receptivity to certain genres of Spanish language music (based on anecdotal observation).

Regarding English media exposure, the results depict an even larger generational divergence; despite its ubiquity relative to Spanish language media, the low English language proficiency levels among the G1 appear to influence their consumption habits considerably. Interestingly, the data suggests that the G2 consumes more English language media than does the G1 Spanish language media, implying perhaps that the proportionate consumption of Swedish language media is greater for the G1 (comparable Swedish media consumption data is unfortunately unavailable). In addition, the data in Table 4-7 indicate that the G1’s consumption of English media is greater than the G2’s of Spanish media, despite the G1’s limited proficiency in English relative to the G2’s command of Spanish (see Table 4-8).

Table 4-8 summarizes the self-reported proficiency by both groups in Swedish, Spanish, and English.
TABLE 4-8
SELF-REPORTED PROFICIENCY ACROSS LANGUAGES

<table>
<thead>
<tr>
<th></th>
<th>G1</th>
<th>G2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Swedish</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>3.57</td>
<td>4.91</td>
</tr>
<tr>
<td>Comprehension</td>
<td>3.65</td>
<td>4.92</td>
</tr>
<tr>
<td>Reading</td>
<td>3.78</td>
<td>4.84</td>
</tr>
<tr>
<td>Writing</td>
<td>3.57</td>
<td>4.96</td>
</tr>
<tr>
<td></td>
<td>3.22</td>
<td>4.68</td>
</tr>
<tr>
<td><strong>Spanish</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>4.45</td>
<td>3.16</td>
</tr>
<tr>
<td>Comprehension</td>
<td>4.73</td>
<td>3.83</td>
</tr>
<tr>
<td>Reading</td>
<td>4.45</td>
<td>3.96</td>
</tr>
<tr>
<td>Writing</td>
<td>4.18</td>
<td>3.42</td>
</tr>
<tr>
<td></td>
<td>4.38</td>
<td>2.96</td>
</tr>
<tr>
<td><strong>English</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>1.80</td>
<td>3.78</td>
</tr>
<tr>
<td>Comprehension</td>
<td>1.71</td>
<td>3.58</td>
</tr>
<tr>
<td>Reading</td>
<td>1.71</td>
<td>4.02</td>
</tr>
<tr>
<td>Writing</td>
<td>1.81</td>
<td>3.79</td>
</tr>
<tr>
<td></td>
<td>1.52</td>
<td>3.29</td>
</tr>
</tbody>
</table>

Note: All four skills rated on a 5-point scale.

For the G2, the perceived order of overall dominance was Swedish, English, and, in last place, Spanish. This order does not mirror the speakers’ order of acquisition, but seems to coincide with the speakers’ current patterns of use (Table 4-7). In Swedish, they evaluated most positively their speaking and reading skills. In Spanish, the highest rating was assigned to the skill of comprehension, while writing was rated lowest. In fact, reflecting their limited formal instruction, the literacy-related skills of writing and reading were rated considerably lower than speaking and comprehension. In English, they considered themselves most proficient in the receptive skills of comprehension and reading, with lower ratings assigned to speaking and writing. Such a hierarchy seems to be influenced by the media-heavy nature of Swedish youths’ exposure to English (Sharp 2001), even though they also complete compulsory English study in school.
The G1 self-reports revealed a distinct order of dominance: Spanish, Swedish, English. In Spanish, they rated themselves highest in speaking and lowest in reading, perhaps due to the reduced access over the years to Spanish language literature. In Swedish, the highest rated skill was comprehension, the lowest writing; furthermore, speaking was rated higher than the receptive skill of reading, perhaps due to a rather oral/aural acquisition of the societal language (in the workplace, TV, radio). In English, though they reported very little proficiency, reading was rated most favorably and writing the lowest; this suggests that the G1, too, is most exposed to receptive English input through media sources.

Table 4-9 presents the results of a multiple correlation analysis in which the relationships of the previously discussed language variables with total lexical productivity (i.e., total word production in all 21 domains) are tested. For the G1, performance on the Spanish proficiency measure was positively correlated with total word production ($r = .354, p < .05$). Weaker correlations were also found for self-reported Spanish proficiency and English media consumption ($p < .10$). Thus, for the G1, the proficiency variables were all positive predictors of lexical productivity. Strangely, Spanish media consumption and reported Spanish use were negatively correlated with total word production for this group.
TABLE 4-9
CORRELATIONS BETWEEN LANGUAGE VARIABLES AND TOTAL WORDS

<table>
<thead>
<tr>
<th></th>
<th>Total Words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G1</td>
</tr>
<tr>
<td>HL education</td>
<td></td>
</tr>
<tr>
<td>Spanish proficiency</td>
<td>.354*</td>
</tr>
<tr>
<td>English proficiency</td>
<td></td>
</tr>
<tr>
<td>Spanish use</td>
<td>-.675**</td>
</tr>
<tr>
<td>Spanish media</td>
<td>-.532**</td>
</tr>
<tr>
<td>English media</td>
<td>.318†</td>
</tr>
<tr>
<td>Self-report Spanish</td>
<td>.304†</td>
</tr>
<tr>
<td>Self-report English</td>
<td>.173</td>
</tr>
</tbody>
</table>

Note. Correlations are Pearson coefficients
† p < 0.10, * p < 0.05, ** p < .001

For the G2, amount of HL education was highly positively correlated with total word production \( r = .69, p < .001 \), suggesting that HL instruction benefits lexical development.

Alternatively, this result could also simply indicate that HS with stronger proficiency (i.e., lexical knowledge) and/or more language support in other areas of their lives were more apt to participate in “mother tongue” education, either of their own accord or due to parental wishes. In addition, English proficiency test results \( r = .446, p < .05 \) and self-reported English proficiency \( r = .482, p < .05 \) were significantly positively correlated with total word production, but English media consumption was not. This suggests that early L3 English acquisition does not hinder HL development but instead supports it. An individual’s general aptitude for language learning could also contribute to this result, a relationship meriting a closer look. The relationship between total word production and performance on the Spanish proficiency measure approached significance \( r = .297, p < .10 \), suggesting, as other research has, that lexical knowledge as measured via tests of lexical availability may serve as a reliable measure of global HL proficiency. Other studies (e.g., Fairclough 2011, 2013a) using different combinations of
measures have found stronger correlations, however, motivating a closer look at the relationships between specific types of grammatical skills and aspects of lexical knowledge. No significant relationship was found between word production and Spanish use, Spanish media consumption, or self-reported Spanish proficiency. That is, HS who perceived themselves as using and being exposed to more Spanish in their daily activities, as well as having stronger proficiency, did not produce significantly more words; likewise, those who perceived their use, input, and proficiency as infrequent and weak, did not generate significantly fewer words.

4.3 Research Question 2a: Lexical Compatibility

The following sections address the research questions pertaining to the quality of the lexical data. In order to provide analyses of appropriate depth, six of the 21 domains were examined. As discussed in Chapter 3, these were selected strategically in order to represent different areas of the basic-abstract (domestic-extradomestic) domain continuum.

Table 4-10 provides a summary of the between-group compatibility tests performed on the groups' collective lexical output via the online program at www.dispolex.com. The subset of items analyzed included only word types with an AI of .02 or higher (representing roughly 80% of all tokens) in each domain. As explained in Chapter 3, this action excluded many late-mentioned, idiosyncratic lexical items among the speakers' productions, permitting a focus on the words most representative of each group's collective vocabulary (Samper Padilla & Samper Hernández 2007). The domains are ranked below according to their degree of compatibility, i.e., the extent to which the two groups' words coincided.
TABLE 4-10  
OVERVIEW OF INTERGENERATIONAL LEXICAL COMPATIBILITY  
IN SIX DOMAINS

<table>
<thead>
<tr>
<th>Rank</th>
<th>Domain</th>
<th>Compatibility (%)</th>
<th>Shared G2 Word Types</th>
<th>Shared G1 Word Types</th>
<th>G2 % Words Analyzed</th>
<th>G1 % Words Analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Human Body</td>
<td>39.24</td>
<td>62</td>
<td>104</td>
<td>68</td>
<td>68</td>
</tr>
<tr>
<td>3</td>
<td>Furniture</td>
<td>17.39</td>
<td>16</td>
<td>34</td>
<td>36</td>
<td>72</td>
</tr>
<tr>
<td>2</td>
<td>Modes of Transportation</td>
<td>23.81</td>
<td>20</td>
<td>38</td>
<td>34</td>
<td>76</td>
</tr>
<tr>
<td>6</td>
<td>Professions and Trades</td>
<td>10.39</td>
<td>16</td>
<td>50</td>
<td>51</td>
<td>81</td>
</tr>
<tr>
<td>4</td>
<td>Economy</td>
<td>13.92</td>
<td>22</td>
<td>51</td>
<td>61</td>
<td>86</td>
</tr>
<tr>
<td>5</td>
<td>Politics</td>
<td>12.62</td>
<td>26</td>
<td>73</td>
<td>69</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19.56</td>
<td>27</td>
<td>58.3</td>
<td>53.2</td>
<td>79</td>
</tr>
</tbody>
</table>

Note: Analysis includes all word types with an AI of .02 or higher in the group lists.

The G2 subset included an average of 58.33 word types per domain, while the G1 average was slightly lower at 53.17, showing once again the G2's greater response variability. The rightmost columns display the proportion of G2 and G1 words included in the analysis. For example, in Furniture, the 34 and 36 words analyzed represented 72% of the G2's and 64% of the G1's total word types.

Results show an overall lexical compatibility of 19.56%. Human Body achieved the highest compatibility ranking by far with a 39.24% (62 words) overlap between the groups. Next in ranking was Modes of Transportation with a 23.81% overlap. This was followed by Furniture, for which the groups' words coincided at a rate of 17.39%. The top ranking of these first three domains with respect to compatibility is in line with expectations that their more basic, concrete nature would favor similar interpretations and experiences (and thus lexicon) in these areas. As for the remaining three domains, it is interesting that Professions and Trades, a seemingly somewhat basic category, is ranked last, even behind The Economy and Politics. Not only did the groups produce statistically significantly different word quantities in this domain, as seen earlier
in the chapter, but now we also see that their words were quite distinct from one another (10.39% compatibility). Ahead, analyses of the actual words upon which these comparisons are based help to identify some of the patterns responsible for such differences, and, by extension, the areas meriting incorporation into classroom instruction.

It was also of interest to measure the degree to which speakers of the same group responded alike. Table 4-11 provides a direct measure of each group's lexical diversity across the six domains. For this the Guiraud’s Index (Root TTR) was employed. It was chosen for its ability to provide reliable comparisons across texts of different sizes (the number of responses varied across domains). It consists of dividing the number of word types by the square root of the number of tokens in a given text (in this study, the collective group word lists). Higher values signal greater diversity among the participant responses and, in lexical availability studies, themes that are open to new vocabulary. Conversely, lower values signal less response uniformity, and concentrated, closed categories. These results revealed greater lexical diversity on the part of both groups toward the end of the continuum, where domain topics represent semantic areas less core to daily life. These data show that overall the G1 and G2 exhibit similar patterns of lexical diversity across the six domains examined.

<table>
<thead>
<tr>
<th>TABLE 4-11</th>
<th>GIRAUDS INDEXES IN COMPARISON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>G2</td>
</tr>
<tr>
<td>The Human Body</td>
<td>4.8</td>
</tr>
<tr>
<td>Furniture</td>
<td>2.1</td>
</tr>
<tr>
<td>Modes of Transportation</td>
<td>2.2</td>
</tr>
<tr>
<td>Professions and Trades</td>
<td>2.9</td>
</tr>
<tr>
<td>The Economy</td>
<td>3.1</td>
</tr>
<tr>
<td>Politics</td>
<td>4.5</td>
</tr>
<tr>
<td>Total</td>
<td>3.3</td>
</tr>
</tbody>
</table>
However, in Table 4-11 we also observe consistently higher lexical diversity indexes for the G2. An extreme degree of response heterogeneity not only signals a low degree of shared experiences among speakers of a group, but can also signal proficiency deficiencies and lack of experience with the language. While the results are certainly indicative of the G2 speakers' robust word knowledge, their consistently greater response variability within the context of a defined semantic domain may also denote “less well-organized lexicons” (Schmitt 2010, p. 18).

In his lexical research among native English speakers, Schmitt (2010) noted that their “association responses exhibit a great deal of systematicity…signalling similar lexical organization.” Conversely, he found that non-native English speakers produced a wider variety of words than natives, suggesting less developed (i.e., less systematic) networks in the mental lexicon. Thus, ironically, the high degree of G2 lexical diversity in relation to that of the G1 may in fact be a function of truncated development in the way of underdeveloped networks in the mental lexicon.

To appreciate some of the concrete differences contributing to the patterns summarized in Tables 4-10 and 4-11, the following sub-sections (4.3.1 to 4.3.4) are dedicated to a qualitative analyses of the parallel and divergent between-group vocabulary in four of the domains (Human Body, Furniture, Modes of Transportation, and Politics). The analysis focuses only on these four due to spacial constraints, and also because I believe they sufficiently highlight recurrent patterns. Later in the Chapter (Section 4.6), some perspectives will be presented regarding the groups’ distinct conceptualizations of the six domain themes, which will provide additional insight into the between-group response discontinuities.

19 For example, in their large-scale Chilean corpus, Valencia & Echeverría (1999) found the type/token ratio to vary significantly between rural and urban youth, with the rural participants’ more homogeneous responses revealing a higher degree of shared experiences among members of that group. The G2's greater lexical diversity displayed in Table 4-11 support what we know about the especially variable linguistic, cultural, and educational experiences of the second generation vis-à-vis their parent generation.
4.3.1 Domain one: Human Body

This and the following sub-sections compare and contrast the Top 50 words of the groups’ AI output lists. These subsets were used as a sample of the most available words from which patterns of difference could be drawn. Contact lexicon appearing in these lists will be discussed in Section 4.4.

<table>
<thead>
<tr>
<th>Shared</th>
<th>G2 Only</th>
<th>G1 Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Gloss</td>
<td>Item</td>
</tr>
<tr>
<td>mano</td>
<td>hand</td>
<td>labio</td>
</tr>
<tr>
<td>corazón</td>
<td>heart</td>
<td>cabello</td>
</tr>
<tr>
<td>cabeza</td>
<td>head</td>
<td>mejilla</td>
</tr>
<tr>
<td>pie</td>
<td>foot</td>
<td>cadera</td>
</tr>
<tr>
<td>ojo</td>
<td>eye</td>
<td>pecho</td>
</tr>
<tr>
<td>dedo</td>
<td>finger</td>
<td>frente</td>
</tr>
<tr>
<td>nariz</td>
<td>nose</td>
<td>pestaña</td>
</tr>
<tr>
<td>boca</td>
<td>mouth</td>
<td>cara</td>
</tr>
<tr>
<td>pelo</td>
<td>hair</td>
<td>piel</td>
</tr>
<tr>
<td>pierna</td>
<td>leg</td>
<td>trasero</td>
</tr>
<tr>
<td>estómago</td>
<td>stomach</td>
<td>pera</td>
</tr>
<tr>
<td>brazo</td>
<td>arm</td>
<td>guata</td>
</tr>
<tr>
<td>oreja</td>
<td>ear</td>
<td>dolor</td>
</tr>
<tr>
<td>uña</td>
<td>nail</td>
<td>bacteria</td>
</tr>
<tr>
<td>cerebro</td>
<td>brain</td>
<td>infección</td>
</tr>
<tr>
<td>hueso</td>
<td>bone</td>
<td>bendición</td>
</tr>
<tr>
<td>espalda</td>
<td>back</td>
<td>gordo</td>
</tr>
<tr>
<td>pulmón</td>
<td>lung</td>
<td>fuerte</td>
</tr>
<tr>
<td>hombro</td>
<td>shoulder</td>
<td>largo</td>
</tr>
<tr>
<td>músculo</td>
<td>muscle</td>
<td>alto</td>
</tr>
<tr>
<td>rodilla</td>
<td>knee</td>
<td></td>
</tr>
<tr>
<td>sangre</td>
<td>blood</td>
<td></td>
</tr>
<tr>
<td>cuello</td>
<td>neck</td>
<td></td>
</tr>
<tr>
<td>diente</td>
<td>tooth</td>
<td></td>
</tr>
<tr>
<td>lengua</td>
<td>tongue</td>
<td></td>
</tr>
<tr>
<td>ceja</td>
<td>eyebrow</td>
<td></td>
</tr>
<tr>
<td>páncreas</td>
<td>pancreas</td>
<td></td>
</tr>
<tr>
<td>enfermedad</td>
<td>sickness/disease</td>
<td></td>
</tr>
</tbody>
</table>

*Chileanism, or regionalism (Diccionario de americanismos, 2010)

*Swedish borrowing
The ‘Shared’ column of Table 4-12 displays the many words, both anatomically simple and sophisticated, that both groups readily mentioned in *Human Body*: mano ‘hand’, estómago ‘stomach’, rodilla ‘knee’, sangre ‘blood’, pulmón ‘lung’, páncreas ‘pancreas’. Among the items that differ, important to note are the many semantically peripheral nouns: G1 — fuerza ‘strength’, vejez ‘old age’, aroma ‘smell, scent’, vida ‘life’, valor ‘strength’; G2 — bendición ‘blessing’, infección ‘infection’, bacteria ‘bacteria’, dolor ‘pain’. Distinct descriptive adjectives are also observed: G1 — hermoso ‘beautiful’, femenino ‘feminine’, sano ‘healthy’; G2 — largo ‘long, tall’, alto ‘tall’, gordo ‘fat’, fuerte ‘strong’. Frequency effects seem to play somewhat of a role; for example, while the G2 showed a preference for the adjectival variant fuerte ‘strong’, the G1 produced the more infrequent fuerza ‘strength’. That example could also reflect the tendency of less advanced speakers toward syntagmatic rather than paradigmatic word associations (Schmitt 2010).

The word largo ‘long’ in the G2 list (in variation with alto, also listed) appears to be a semantic extension resulting from the Swedish use of lång ‘long’ to describe stature (more on this is the Section 4.4). One verb was produced by the G1: matar ‘to kill’. Other noteworthy contrasts were stylistic in nature, having to do with technical or higher register words versus colloquial, even taboo vocabulary: G1 — cráneo ‘cranium’, seso ‘brain’, esqueleto ‘skeleton’, pantorrilla ‘calf’; G2 — trasero, poto ‘butt’, pera ‘chin’, guata ‘belly’. Figure 4-3 provides a visual representation of some of the commonalities and differences discussed here.

Figure 4-3. Sample G2-G1 divergent/convergent lexicon (*Human Body*)
4.3.2 Domain two: Furniture

Table 4-13 displays the distribution of the vocabulary offered under the theme of Furniture.

Among the coinciding words are many prototypical items such as cama ‘bed’, mesa ‘table’ and silla ‘chair’. Interestingly, both groups also overlapped in the peripheral concepts of alfombra ‘carpet’ and the descriptors moderno ‘modern’, antiguo ‘old, antique’ and café ‘brown’. The groups coincided in the preference for the widespread Anglicism clóset ‘closet’, although the G1 also produced the variants ropero and guardarropa ‘closet’.
TABLE 4-13
WORDS IN COMPARISON (FURNITURE)

<table>
<thead>
<tr>
<th>Item</th>
<th>G2 Only</th>
<th>G1 Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>mesa</td>
<td>mesa</td>
<td>mesa</td>
</tr>
<tr>
<td>silla</td>
<td>sillón</td>
<td>sillón</td>
</tr>
<tr>
<td>sillón</td>
<td>sofá</td>
<td>sofá</td>
</tr>
<tr>
<td>cama</td>
<td>estante (de libro)</td>
<td>estante (de libro)</td>
</tr>
<tr>
<td>banco</td>
<td>comedor</td>
<td>velador</td>
</tr>
<tr>
<td>televisión</td>
<td>velador</td>
<td>velador</td>
</tr>
<tr>
<td>clóset</td>
<td>madera</td>
<td>moderno</td>
</tr>
<tr>
<td>madera</td>
<td>moderno</td>
<td>antiguo</td>
</tr>
<tr>
<td>moderno</td>
<td>café</td>
<td>café</td>
</tr>
<tr>
<td>Item</td>
<td>Gloss</td>
<td>Item</td>
</tr>
<tr>
<td>mesa</td>
<td>table</td>
<td>ropero</td>
</tr>
<tr>
<td>silla</td>
<td>chair</td>
<td>guardarropa</td>
</tr>
<tr>
<td>sillón</td>
<td>armchair</td>
<td>cómoda</td>
</tr>
<tr>
<td>cama</td>
<td>bed</td>
<td>vitrina</td>
</tr>
<tr>
<td>sofá</td>
<td>couch</td>
<td>aparador</td>
</tr>
<tr>
<td>estante (de libro)</td>
<td>bookshelf</td>
<td>bufe</td>
</tr>
<tr>
<td>alfombra</td>
<td>carpet/rug</td>
<td>colgador</td>
</tr>
<tr>
<td>comedor</td>
<td>dining table</td>
<td>cena</td>
</tr>
<tr>
<td>velador</td>
<td>nightstand</td>
<td>cuna</td>
</tr>
<tr>
<td>banco</td>
<td>bench</td>
<td>catre</td>
</tr>
<tr>
<td>televisión</td>
<td>TV</td>
<td>banca</td>
</tr>
<tr>
<td>clóset</td>
<td>closet</td>
<td>bukila</td>
</tr>
<tr>
<td>madera</td>
<td>wood</td>
<td>decoración</td>
</tr>
<tr>
<td>moderno</td>
<td>modern</td>
<td>radio</td>
</tr>
<tr>
<td>antiguo</td>
<td>antique</td>
<td>florero</td>
</tr>
<tr>
<td>café</td>
<td>brown</td>
<td>tetera</td>
</tr>
<tr>
<td>Item</td>
<td>Gloss</td>
<td>Item</td>
</tr>
<tr>
<td>mesa</td>
<td>working table, island</td>
<td>ropero</td>
</tr>
<tr>
<td>silla</td>
<td>dishwasher</td>
<td>guardarropa</td>
</tr>
<tr>
<td>sillón</td>
<td>coffee table</td>
<td>cómoda</td>
</tr>
<tr>
<td>cama</td>
<td>desk</td>
<td>vitrina</td>
</tr>
<tr>
<td>sofá</td>
<td>bookshelf</td>
<td>aparador</td>
</tr>
<tr>
<td>estante (de libro)</td>
<td>stool</td>
<td>bufe</td>
</tr>
<tr>
<td>alfombra</td>
<td>door</td>
<td>colgador</td>
</tr>
<tr>
<td>comedor</td>
<td>lamp</td>
<td>cena</td>
</tr>
<tr>
<td>velador</td>
<td>painting, picture</td>
<td>cuna</td>
</tr>
<tr>
<td>banco</td>
<td>tree</td>
<td>catre</td>
</tr>
<tr>
<td>televisión</td>
<td>mirror</td>
<td>banca</td>
</tr>
<tr>
<td>clóset</td>
<td>watch</td>
<td>bukila</td>
</tr>
<tr>
<td>madera</td>
<td>seat</td>
<td>decoración</td>
</tr>
<tr>
<td>moderno</td>
<td>bedroom</td>
<td>cuadro</td>
</tr>
<tr>
<td>antiguo</td>
<td>sex</td>
<td>radio</td>
</tr>
<tr>
<td>café</td>
<td>money</td>
<td>florero</td>
</tr>
<tr>
<td>Item</td>
<td>Gloss</td>
<td>Item</td>
</tr>
<tr>
<td>mesa</td>
<td>store</td>
<td>tetera</td>
</tr>
<tr>
<td>silla</td>
<td>movie</td>
<td>olla</td>
</tr>
<tr>
<td>sillón</td>
<td>home</td>
<td>cerrojo</td>
</tr>
<tr>
<td>cama</td>
<td>bed pillow</td>
<td>boutique</td>
</tr>
<tr>
<td>sofá</td>
<td>decoration</td>
<td>estilo</td>
</tr>
<tr>
<td>estante (de libro)</td>
<td>party</td>
<td>masaje</td>
</tr>
<tr>
<td>alfombra</td>
<td>party</td>
<td>living</td>
</tr>
<tr>
<td>comedor</td>
<td>layer of dust</td>
<td>cocina</td>
</tr>
<tr>
<td>velador</td>
<td>expensive</td>
<td>terraza</td>
</tr>
<tr>
<td>banco</td>
<td>beige</td>
<td>barnizado</td>
</tr>
<tr>
<td>televisión</td>
<td>comfortable</td>
<td>bonito</td>
</tr>
<tr>
<td>clóset</td>
<td>poor</td>
<td>grande</td>
</tr>
<tr>
<td>madera</td>
<td>of-fabric</td>
<td>barato</td>
</tr>
<tr>
<td>moderno</td>
<td>dark color</td>
<td>nuevo</td>
</tr>
<tr>
<td>antiguo</td>
<td>to sell</td>
<td>diferente</td>
</tr>
<tr>
<td>café</td>
<td></td>
<td>blanco</td>
</tr>
<tr>
<td>Item</td>
<td>Gloss</td>
<td>Item</td>
</tr>
<tr>
<td>mesa</td>
<td></td>
<td>espacioso</td>
</tr>
<tr>
<td>silla</td>
<td></td>
<td>de organizar</td>
</tr>
<tr>
<td>sillón</td>
<td></td>
<td>dormir</td>
</tr>
</tbody>
</table>

*Chileanism, or regionalism typical in Chile (Diccionario de americanismos, 2010)

bSwedish borrowing

‘Literally ‘small table’. Can function as a shortened form of mesita de noche ‘nightstand’, mesita de centro ‘coffee table’, or mesita auxiliar ‘end table’.

The divergent vocabulary in this domain consisted of many adjectives, especially abundant and descriptive in the G1 list (e.g., bonito ‘pretty’, barnizado ‘varnished’, espacioso ‘spacious’). While the groups coincided in estante (de libro) ‘(book)shelf’, there were variants exclusive to one group or the other: G1 — bukila (Swedish loanword); G2 — librera. Both
groups supplied additional furniture staples, though the G2 fewer: e.g., G1 — cómoda ‘dresser’, colgador ‘coat rack’, cuna ‘crib’, vitrina ‘glass cabinet’; G2 — mesita ‘nightstand, escritorio ‘desk’, piso ‘stool’. Thus the G2 list contained a larger proportion of semantically peripheral words: e.g., adorno ‘decoration’, tavla (Swedish loanword) ‘picture’, película ‘movie’, hogar ‘home’.

Additional observations included the G2’s mention of taboo words like sexo ‘sex’ (also observed in Human Body), as well as seemingly semantically extended terms like almohada ‘bed pillow’ instead of cojín ‘couch pillow’ or árbol ‘tree’ to refer to planta de interior ‘houseplant’.

What is abundantly clear from the data, however, is that the HS possess an extensive, accessible vocabulary on the topic, even producing more standard terms like dormitorio ‘bedroom’ at a greater rate than highly regional ones like pieza ‘bedroom’ (which, with a much lower AI, appears beyond the Top 50).

Figure 4-4. Sample G2-G1 divergent/convergent lexicon (Furniture)
### 4.3.3 Domain three: Modes of Transportation

Table 4-14 presents the data corresponding to Modes of Transportation.

#### TABLE 4-14
WORDS IN COMPARISON (MODES OF TRANSPORTATION)

<table>
<thead>
<tr>
<th>Shared</th>
<th>Gloss</th>
<th>Item</th>
<th>Gloss</th>
<th>G2 Only</th>
<th>Item</th>
<th>Gloss</th>
<th>G1 Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>auto</td>
<td>car&quot;</td>
<td>Item</td>
<td>Gloss</td>
<td>barco</td>
<td>Item</td>
<td>Gloss</td>
<td>Item</td>
</tr>
<tr>
<td>bus</td>
<td>colectivo</td>
<td>car</td>
<td>shared taxi</td>
<td>crucero</td>
<td>boat</td>
<td>boat</td>
<td></td>
</tr>
<tr>
<td>tren</td>
<td>coche deportivo</td>
<td>boat</td>
<td>sports car</td>
<td>metro</td>
<td>van</td>
<td>van</td>
<td></td>
</tr>
<tr>
<td>avión</td>
<td>lancha</td>
<td>vehicle</td>
<td>boat</td>
<td>tunelbana</td>
<td>caballo</td>
<td>van</td>
<td></td>
</tr>
<tr>
<td>coche</td>
<td>vehículo</td>
<td>scooter</td>
<td>minibus</td>
<td>camello</td>
<td>burro</td>
<td>horse</td>
<td></td>
</tr>
<tr>
<td>bicicleta</td>
<td>patineta</td>
<td>skateboard</td>
<td>yacht</td>
<td>muleta</td>
<td>piernas</td>
<td>donkey</td>
<td></td>
</tr>
<tr>
<td>taxi</td>
<td>minibús</td>
<td>cargo boat</td>
<td>stairs</td>
<td>camioneta</td>
<td>motoneta</td>
<td>legs</td>
<td></td>
</tr>
<tr>
<td>autobús</td>
<td>monopatín</td>
<td>wings</td>
<td>walk, stroll</td>
<td>camello</td>
<td>carreta</td>
<td>motor scooter</td>
<td></td>
</tr>
<tr>
<td>moto</td>
<td>barco de cargo</td>
<td>walk</td>
<td>fly</td>
<td>muleta</td>
<td>carreté</td>
<td>wagon, cart</td>
<td></td>
</tr>
<tr>
<td>micro</td>
<td>escalera</td>
<td>cargo</td>
<td>run</td>
<td>bicimoto</td>
<td>minibús</td>
<td>van, truck</td>
<td></td>
</tr>
<tr>
<td>bote</td>
<td>alas</td>
<td>wheel</td>
<td>cargo</td>
<td>flotador</td>
<td>minibus</td>
<td>camel</td>
<td></td>
</tr>
<tr>
<td>camión</td>
<td>andar</td>
<td>gas</td>
<td>wheel</td>
<td>tranvia</td>
<td>bicycle</td>
<td>crush</td>
<td></td>
</tr>
<tr>
<td>automóvil</td>
<td>volar</td>
<td>driver</td>
<td>pilot</td>
<td>motocicleta</td>
<td>camión</td>
<td>motorbike</td>
<td></td>
</tr>
<tr>
<td>patín</td>
<td>caminar</td>
<td>steer</td>
<td>wheel</td>
<td>patín de hielo</td>
<td>viajar</td>
<td>trolley</td>
<td></td>
</tr>
<tr>
<td></td>
<td>correr</td>
<td>gas</td>
<td>driver</td>
<td>pagar</td>
<td>pago</td>
<td>motorcycle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cargo</td>
<td>wheel</td>
<td>driver</td>
<td>viaje</td>
<td>viaje</td>
<td>iceskate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rueda</td>
<td>wheel</td>
<td>pilot</td>
<td>petróleo</td>
<td>trip</td>
<td>travel</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bencina</td>
<td>gas</td>
<td>chofer, driver</td>
<td>pasajero</td>
<td>gas, petrol</td>
<td>pay</td>
<td></td>
</tr>
<tr>
<td></td>
<td>conductor</td>
<td>gas</td>
<td>chofer, driver</td>
<td>accidente</td>
<td>passenger</td>
<td>trip</td>
<td></td>
</tr>
<tr>
<td></td>
<td>piloto</td>
<td>gas</td>
<td>mechanical</td>
<td>nausea</td>
<td>accident</td>
<td>accident</td>
<td></td>
</tr>
<tr>
<td></td>
<td>chofer</td>
<td>pilot</td>
<td>mechanic</td>
<td>por mar</td>
<td>nausea</td>
<td>nausea</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mecánico</td>
<td>pilot</td>
<td>station</td>
<td>por tierra</td>
<td>by sea</td>
<td>by sea</td>
<td></td>
</tr>
<tr>
<td></td>
<td>estación</td>
<td>speed</td>
<td>station</td>
<td>or la</td>
<td>by land</td>
<td>by land</td>
<td></td>
</tr>
<tr>
<td></td>
<td>velocidad</td>
<td>safety</td>
<td>speed</td>
<td>olor</td>
<td>smell, odor</td>
<td>smell, odor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>seguridad</td>
<td>hurry</td>
<td>safety</td>
<td>trabajo</td>
<td>heat</td>
<td>heat</td>
<td></td>
</tr>
<tr>
<td></td>
<td>prisa</td>
<td>line</td>
<td>hurry</td>
<td>necesidad</td>
<td>necessity</td>
<td>necessity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>línea</td>
<td>passport</td>
<td>line</td>
<td>ruido</td>
<td>work</td>
<td>work</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pasaporte</td>
<td>studies</td>
<td>passport</td>
<td>calor</td>
<td>noise</td>
<td>noise</td>
<td></td>
</tr>
<tr>
<td></td>
<td>estudios</td>
<td>country</td>
<td>studies</td>
<td>niños</td>
<td>heat</td>
<td>heat</td>
<td></td>
</tr>
<tr>
<td></td>
<td>país</td>
<td>country</td>
<td>country</td>
<td>verdura</td>
<td>children</td>
<td>children</td>
<td></td>
</tr>
<tr>
<td></td>
<td>susto</td>
<td>country</td>
<td>country</td>
<td>fácil</td>
<td>vegetables</td>
<td>vegetables</td>
<td></td>
</tr>
<tr>
<td></td>
<td>construcción</td>
<td>construction</td>
<td>stress</td>
<td>necesario</td>
<td>easy</td>
<td>easy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>stress</td>
<td>construction</td>
<td>stress</td>
<td>mucho</td>
<td>necessary</td>
<td>necessary</td>
<td></td>
</tr>
<tr>
<td></td>
<td>experiment</td>
<td>experiment</td>
<td>experiment</td>
<td>dumber</td>
<td>crowd</td>
<td>crowd</td>
<td></td>
</tr>
</tbody>
</table>

*Chileanism, or regionalism typical in Chile (Diccionario de americanismos, 2010)

*Swedish borrowing
The marginal amount of overlap among the top words in this domain included the most prototypical concepts related to the topic, including both the Southern Cone variant of *auto(móvil)* ‘car’ and the Peninsular *coche* ‘car’. In addition, the G2 list included *carro* ‘car’, which suggests exposure to other Latin American dialects of Spanish. The G2 also stood out in their plentiful production of various infinitival forms: *caminar* ‘to walk’, *andar* ‘to walk, to ride’, *correr* ‘to run’, *volar* ‘to fly’. Again, however, the G1 adhered more closely to the semantic theme through their greater production of actual means of transportation: e.g., *barco* ‘ship’, *metro* ‘subway’, *muleta* ‘crutch’, *bicimoto* ‘motorbike’. By contrast, the G2 relied more heavily on words only related to means of transportation, perhaps constituting the use of circumlocutive strategies in the face of lexical gaps: e.g., *rueda* ‘wheel’, *bencina* ‘gas’, *piloto* ‘pilot’, *velocidad* ‘speed’. Swedish and English borrowings, to be fully discussed in Section 4.4, figured among the group’s individual lists, though none was found among the shared items. Figure 4-5 provides a sample of the groups' intersecting and divergent lexicon.


4.3.4 Domain four: Politics

In the domain of Politics (Table 4-15), the groups’ shared productions included many basic titles and concepts, such as ministro 'minister', político 'politician', izquierda 'left' and ley 'law', as well as strongly negatively connoting items such as mentira 'lie' and corrupción 'corruption'. Among the groups’ distinctive vocabulary items, one observed pattern was the tendency of the G2 to produce words denoting concepts (democracia 'democracy', comunismo 'communism', anarquía 'anarchy', república 'republic') while the G1 preferred mentioning the adherents to such concepts (demócrata 'democrat', comunista 'communist', anarquista 'anarchist', gobernador 'governor'). A related thematic observation was the G2’s mention of more extreme affiliations such as nazista 'nazi', fascista 'fascist', and dicatador/dictadura 'dictator/dictatorship'. In fact, the G2’s list appears to reflect an overall more negative conception of the theme with words like puerco 'pig',
violencia 'violence', amargo 'bitter', complicado 'complicated' and odio 'hate', though G1 items like sucio 'dirty', mentiroso 'liar', estafa 'scam, con', and difícil 'difficult' point to similar critiques.

### TABLE 4-15
WORDS IN COMPARISON (POLITICS)

<table>
<thead>
<tr>
<th>Shared</th>
<th>G2 Only</th>
<th>G1 Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Gloss</td>
<td>Item</td>
</tr>
<tr>
<td>presidente</td>
<td>presidente</td>
<td>dinero</td>
</tr>
<tr>
<td>problema</td>
<td>problema</td>
<td>país</td>
</tr>
<tr>
<td>gobierno</td>
<td>gobierno</td>
<td>elección</td>
</tr>
<tr>
<td>derecha</td>
<td>derecha</td>
<td>democracia</td>
</tr>
<tr>
<td>partido</td>
<td>partido</td>
<td>derechos</td>
</tr>
<tr>
<td>izquierda</td>
<td>izquierda</td>
<td>plata</td>
</tr>
<tr>
<td>ministro</td>
<td>ministro</td>
<td>anarquía</td>
</tr>
<tr>
<td>mentira</td>
<td>mentira</td>
<td>televisión</td>
</tr>
<tr>
<td>ley</td>
<td>ley</td>
<td>complicado</td>
</tr>
<tr>
<td>político</td>
<td>político</td>
<td>nazi</td>
</tr>
<tr>
<td>corrupción</td>
<td>corrupción</td>
<td>revolución</td>
</tr>
<tr>
<td>voto</td>
<td>voto</td>
<td>fome</td>
</tr>
<tr>
<td>poder</td>
<td>poder</td>
<td>horrible</td>
</tr>
<tr>
<td>diario</td>
<td>diario</td>
<td>distinto</td>
</tr>
<tr>
<td>reunión</td>
<td>reunión</td>
<td>robar</td>
</tr>
<tr>
<td>socialista</td>
<td>socialista</td>
<td>puerco</td>
</tr>
<tr>
<td>desacuerdo</td>
<td>desacuerdo</td>
<td>la casa blanca</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rojo</td>
</tr>
<tr>
<td>ideología</td>
<td>ideología</td>
<td>ideología</td>
</tr>
<tr>
<td>comunismo</td>
<td>comunismo</td>
<td>communism</td>
</tr>
<tr>
<td>república</td>
<td>república</td>
<td>republic</td>
</tr>
<tr>
<td>noticia</td>
<td>noticia</td>
<td>news</td>
</tr>
<tr>
<td>fascista</td>
<td>fascista</td>
<td>fascist</td>
</tr>
<tr>
<td>dictador</td>
<td>dictador</td>
<td>dictator</td>
</tr>
<tr>
<td>popular</td>
<td>popular</td>
<td>popular</td>
</tr>
<tr>
<td>tristeza</td>
<td>tristeza</td>
<td>sadness</td>
</tr>
<tr>
<td>tipo de pensar</td>
<td>tipo de pensar</td>
<td>way of thinking</td>
</tr>
<tr>
<td>amargo</td>
<td>amargo</td>
<td>bitter</td>
</tr>
<tr>
<td>socialismo</td>
<td>socialismo</td>
<td>socialism</td>
</tr>
<tr>
<td>falso</td>
<td>falso</td>
<td>fake</td>
</tr>
<tr>
<td>odio</td>
<td>odio</td>
<td>hate</td>
</tr>
<tr>
<td>esposas</td>
<td>esposas</td>
<td>handcuffs</td>
</tr>
<tr>
<td>dictadura</td>
<td>dictadura</td>
<td>dictatorship</td>
</tr>
<tr>
<td>violencia</td>
<td>violencia</td>
<td>violence</td>
</tr>
</tbody>
</table>

Figure 4-6 displays a sample of the shared and unique lexical items pertaining to Politics.
The general impression in the domain of Politics is that the G2 productions are relevant and often quite sophisticated. A few differences were observed, however. With respect to synonymic and stylistic variety, while the G2 list only included the action robar 'to steal', the G1 list contained the following related nouns: ladrón 'thief', estafa 'scam, con', robo 'robbery'. In addition, once again, the G2 list contained colloquial or regional terms like fome 'boring, stupid' and plata 'cash, dough', while the G1 semantically related counterparts were the less marked, lower frequency terms aburrido 'boring' and riqueza/pobreza 'wealth/poverty'. Lastly, as seen in Human Body, the more frequent mention of adjectives (e.g., complicado 'complicated', horrible 'horrible', distinto 'different', falso 'fake') on the part of the G2 may point to their greater reliance on syntagmatic word associations, which is correlated with less developed mental lexicons (Schmitt 2010).

The data presented in the previous sections have provided a window into the complexity and diversity of the G2 lexical store. Not only were the G2 productions numerically comparable
to those of the G1 across domains, but they also reflect robust lexical knowledge far beyond the most basic and personal domains. The principal intergenerational differences outlined in the previous sections, however, included the G2’s greater production of semantically peripheral items (perhaps reflecting circumlocution) as well as colloquial vocabulary. Their responses also suggested a higher rate of syntagmatically motivated associations, which are characteristic of less semantically developed lexical networks.

4.4 Research Questions 2b: Contact Lexicon

4.4.1 English contact lexicon

I will first address the English-origin lexicon identified in the participant lists, summarized in Table 4-16. Sparse in general, the majority of items were classified as pure borrowings or loanblends, which are closely related categories. The overall density of English borrowings in the G2 dataset was a mere 0.7%, and 0.5% for the G1. In terms of their vitality in the community, the results revealed an average of only 0.5 borrowings per speaker.

Included in those figures are five items that have some degree of extension in general and/or Chilean Spanish, and which thus may qualify as “covert Anglicisms” (Gottlieb 2005), or words of English origin that are not obvious to the speaker. These include closet, living, barman, mitin, and gásfiter, all of which are found in Spanish language dictionaries, including that of the Real Academia Española (RAE), and two of which appear to be of especially wide usage (i.e., the first two). Though these five have almost certainly not emerged as a result of English contact in Sweden, it may be that English exposure has caused an increase in their usage (as opposed to other variants, such as ropero or armario ‘closet’, or sala de estar ‘living room’. In fact, these and other variants were found among the participant productions, though with much lower
frequencies. If not an increase in their usage, it is possible that knowledge of English leads to a heightened perception of their English origin. For these reasons I have included them in the analysis.

TABLE 4-16
RATE AND DISTRIBUTION OF ENGLISH CONTACT LEXICON
<table>
<thead>
<tr>
<th>Category</th>
<th>Human Body</th>
<th>Furniture</th>
<th>Modes of Transport</th>
<th>Professions &amp; Trades</th>
<th>Economy</th>
<th>Politics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>0.6 (4)</td>
<td>0.9 (3)</td>
<td>0.7 (12)</td>
<td>0.7 (4)</td>
<td>0.5 (3)</td>
<td>0.3 (2)</td>
</tr>
<tr>
<td># tokens</td>
<td>12/1899</td>
<td>37/148</td>
<td>27/12</td>
<td>21/67</td>
<td>12/32</td>
<td>13/22</td>
</tr>
<tr>
<td>Density</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Notes</td>
<td>a Not meant for between-group comparison, given unequal n’s. Only important in relation to Totals and Density.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notes</td>
<td>b Number of tokens, if greater than one.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The prediction that the more advanced themes would house a greater proportion of contact lexicon than the basic ones was not confirmed in either the G1 or the G2 data. However, considering only the spontaneous (non-established) borrowings, the prediction obtains. The prediction that the G2 data would show a greater density of English borrowings was confirmed, though the difference was insignificant (0.7% versus 0.5%). However, once again, if the established Anglicisms were to be omitted, this difference would be more marked. The prediction that the G2 would have a greater density of English origin lexicon than the G1 in the basic domains was confirmed by the data, a result that would not be altered with the omission of established borrowings. Notably, there was not a single English borrowing in the domain of Human Body, which, given that only in advanced stages of language shift do borrowings regularly appear in the most basic domains (Nagy 2011), speaks to the G2’s strong bilingualism. Furthermore, the borrowings were predominantly nonce rather than established, attesting to their limited extension and thus vitality in the community.

An additional observation relates to the loanword van, which is also of common use in Swedish, despite the existence of a so-called equivalent Swedish term (skåpbil). This represents an example of how English can reach these speakers indirectly through Swedish. In other words, not only are the HS exposed directly to English discourse through a variety of mediums, they also receive input through the ubiquity of English terms and expressions in everyday Swedish. Given the deluge of English in the speakers’ environment, it is surprising that there were so few Anglicisms included in their lists. Another observation, which coincides with what Moreno-Fernández (2007) noted in his data, was that in some cases participants offered the English and Spanish doublet for a term, attesting to their knowledge of the Spanish form while for some
reason supplying the English equivalent as well. For example, one individual produced both oil and petróleo in close proximity to one another on the list.

4.4.2 Swedish contact lexicon

Table 4-17 displays the Swedish contact lexicon of each domain according to borrowing type. A cursory look reveals that Swedish contact lexicon was more abundant in both data sets than the English contact lexicon. This confirms the clearly dominant role of Swedish relative to English in the linguistic systems and experiences of the speakers. Despite its significant societal status and presence, as well as the G2’s early acquisition of it, English bears much less influence than Swedish on the lexicon of the HL. Nevertheless, in terms of the vitality of Swedish borrowings in the speech community (both generations included), a meager 0.8 borrowings were produced per speaker. Only 1.5% of the G2 tokens, and 0.6% of the G1 tokens, in all six domains constituted Swedish borrowings. These results are considerably more conservative than those of Moreno-Fernández 2007), who reported a borrowing density of between 6.5% and 15% for U.S. SHS across 16 domains. This is perhaps due to his inclusion of domains such as Food and Drink and Clothing in which loanwords may be more likely to appear. In addition, the speakers' access to and membership in a larger, more concentrated Spanish speaking community (Chicago) in Moreno-Fernández's study may contribute to the spread of borrowings.

The domain in which both speaker groups produced the largest proportion of Swedish borrowings was Professions and Trades, followed by Economy (Economy and Politics for the G1). The smallest proportions were found in Human Body (for the G1) and Modes of Transportation (for the G2). However, Professions and Trades was also the domain in which the generations exhibited the largest between-group disparity (4.5% borrowings for the G2 versus 1.3% for the G1), mostly owing to the G2’s many loanblends and phonological intrusions. Only
in *Modes of Transportation* and *Politics* did the G1 density exceed the G2’s, thanks in part to some very interesting loanblends (*idioter* 'idiots', *bablería* 'babble'). Although for both groups, most borrowings were produced under *Professions and Trades*, they followed different patterns. For example, the G1’s borrowings tended to be related to high frequency professions *polis* 'policeman' that they would tend to hear and see often, as well as professions that did not exist or were in their infancy when they left Chile in the 70s, 80s and even 90s (*data-técnico* 'computer technician'). Contrastively, the G2's borrowings in this domain were in large part due to lexical gaps (*jornalista* 'journalist') or phonological intrusions (*fotbolista* 'soccer player').

Overall, the results confirm the prediction that, for both groups, the basic domains would reflect a lower borrowing density than the more complex ones. The prediction that the G2, who are at a more advanced stage of language contact than the G1, would produce a larger proportion of borrowings in the basic domains than would the G1 was also confirmed by the data, though the difference was very slight (1.14% vs. 1%). As a whole, these findings provide evidence against the idea that SHS lexicon bears extensive influence from the majority language.

TABLE 4-17
RATE AND DISTRIBUTION OF SWEDISH CONTACT LEXICON
<table>
<thead>
<tr>
<th>Category</th>
<th>Pure borrowings</th>
<th>Calques</th>
<th>Semantic extensions</th>
<th>Loanblends</th>
<th>Phonological intrusions</th>
<th>Totals</th>
<th>Total Tokens</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Body</td>
<td><em>dna (adn)'dna'</em></td>
<td></td>
<td>largo 'tall'</td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>Furniture</td>
<td><em>(3) tavla (cuadro)</em> 'picture'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
<td>1.1</td>
</tr>
<tr>
<td>Modes of Transportation</td>
<td><em>(2) tunelbana (metro)</em> 'subway'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>Professions and Trades</td>
<td><em>krav (exigencias)</em> 'demands'</td>
<td></td>
<td></td>
<td></td>
<td><em>quemista (químico)</em> [kemist] 'chemist' *sociónomo (asistente social) [socionom] 'social worker' *patrulante (patrullero) [patrullering] 'patrolman' *jornalista (periodista) [journalist] 'journalist' *eléctrico (electricista) [elektriker] 'electrician' *laboratista (laboratorista) [laboratist] 'laboratory technician' <em>officcar (oficial) [officer] 'officer'</em></td>
<td>13</td>
<td>469</td>
<td>4.5</td>
</tr>
<tr>
<td>Economy</td>
<td><em>(2) renta (interés)</em> [ránta] 'interest'</td>
<td></td>
<td><em>en konkurso (a la quiebra)</em> [i konkurs] 'into bankruptcy' *extra trabajo (trabajo extra/adicional) [extra arbete] 'extra work'</td>
<td>*(2) invester (investir) [investera] 'to invest' *ecónomo (economista) [ekonom] 'economist' *cassafiuerte (caja fuerte, seguro) [kassaskåp] 'cashbox', 'safe'</td>
<td>8</td>
<td>279</td>
<td>2.9 0.8</td>
<td></td>
</tr>
<tr>
<td>Politics</td>
<td><em>sammarbete (colaboración)</em> 'collaboration'</td>
<td></td>
<td>*social demócrata (demócrata social) [socialdemokratak] 'social democrat'</td>
<td>*idiôter (idiotas) [idioter] 'idiots' *bablería (balbuceo) [babbel] 'babble, babbling'</td>
<td>1</td>
<td>268</td>
<td>0.3% 0.8%</td>
<td></td>
</tr>
<tr>
<td>Totals (corpus density)</td>
<td><strong>11</strong></td>
<td>7</td>
<td><strong>0</strong></td>
<td><strong>1</strong></td>
<td><strong>11</strong></td>
<td><strong>4</strong></td>
<td><strong>42 (27, 15)</strong></td>
<td><strong>1859</strong> 1.5% 0.6%</td>
</tr>
<tr>
<td>Community density</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.08 per speaker (1.2, 0.5)</td>
</tr>
</tbody>
</table>


aNot meant for between-group comparison, given unequal n’s. Only important in relation to Totals and Density.

bNumber of tokens, if greater than one.
In terms of borrowing types, the results showing a predominance of pure loanwords and loanblends are in line with previous research. While calques were surprisingly only found in the G1 subset, phonological intrusions were exclusive to G2 productions, reflecting the G2’s limited literacy in the HL (that is, their limited experience in viewing many words in their written form). In terms of content, nonce (lone) and established (offered by more than one individual) borrowings were both represented cross-generationally, although a greater portion of the G2 borrowings were established while the G1 produced more nonce borrowings. Informal observations of casual speech produced by the G2, however, suggest that several of the G1 nonce borrowings are also widely used by the G2 in Spanish, such as tunelbana (metro) [tunnelbana] ‘subway’ and social demócrata (demócrata social) [socialdemokrat] ‘social democrat’, even though they did not produce them during the experiment. These findings confirm the existence of well-integrated Swedish-origin lexical items in the Spanish variety of Chilean Swedes, some of which (e.g., data ‘computer’), anecdotally, can even be found in the Spanish lyrics of local hip hop groups composed of G2 ethnic Chileans such as Advance Patrol and The Latin Kings.

Because the English versions of some of the terms in Table 4-17 (dna, quemista 'chemist', budget) are identical or similar to the Swedish forms, it was difficult to determine their original source. Given that Swedish is the majority language and the one of greater command among the participants, I chose to classify them as Swedish borrowings. The points of interaction and overlap between Swedish and English leading to joint lexical effects exerted upon the HL (including cases like van, an Anglicism of wide usage in Swedish) is a topic meriting further study.

Finally, there were several notable orthographic observations. While some of the adapted pure loanwords reflected an elimination of the Swedish consonant gemination (tunelbana
[tunnelbana] ‘subway’), other semi-adapted ones preserved and even amplified it (geminating the c): officcar [officer] ‘officer’. This, along with the participant’s use of the letter c before an /a/ to represent /s/, suggests that the G1 may experience attrition at the orthographic level. The G1’s extensive use of k for /k/ was also noted in the corpus (arkitecto [arquitecto] ‘arquitect’, krédito [crédito] ‘credit’, korona [corona] ‘crown’), including in variation with the G2 (e.g., corona), who seemed rather aware of its particular ‘Swedishness’. However, to confirm attrition of this sort, one would have to assume they had normative orthography prior to arrival in Sweden. The evidence is strong, however, particularly in the extensive use of k, which is unusual in Spanish orthography.

Another curious case involved the G2’s phonologically motivated hybridization of the referentially similar (both meaning ‘safe’), yet semantically distinct [kassaskåp] (lit: ‘cashbox’) and caja fuerte (lit: ‘box strong’). Though phonologically similar, the Swedish [kassa] means ‘cash’, while the Spanish caja signifies ‘box’. The speaker’s fusion of the terms resulted in cassafuerte, literally meaning ‘cash-strong’. Modeling after the term caja fuerte, it appears the speaker reanalyzed caja as ‘cash’ based on its phonological similarity with and identical position of [kassa] in the Swedish compound term.

4.5 Research Question 2c: Dialect Contact (Peninsularisms)

A total of 27 tokens of three Peninsularisms were identified in three of the six domains. They are displayed in Table 4-18.
TABLE 4-18
PENINSULARISMS

<table>
<thead>
<tr>
<th>Peninsularism</th>
<th>Chilean variant</th>
<th>English gloss</th>
<th>G2 (n=23)</th>
<th>G1 (n=30)</th>
<th>Valencia &amp; Echeverría (1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td>coche</td>
<td>auto</td>
<td>car</td>
<td>9</td>
<td>6</td>
<td>✓ (AI: 48)</td>
</tr>
<tr>
<td>ordenador</td>
<td>computador(a)</td>
<td>computer</td>
<td>5</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>mesita</td>
<td>esquinero, mesa de centro</td>
<td>table</td>
<td>2</td>
<td>2</td>
<td>✓ (AI:140)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16 (0.7 per speaker)</td>
<td>11 (0.4 per speaker)</td>
<td></td>
</tr>
</tbody>
</table>

The first two, *coche* 'car' and *ordenador* 'computer'\(^{20}\), are high frequency words in everyday spoken language in Spain, making them likely candidates for borrowing. *Mesita* appears to be a lexicalized, commonly used form in Peninsular Spanish to refer to various kinds of small tables (end table, night stand), whereas Chilean Spanish prefers other terms (*esquinero* 'end table', *velador* 'night stand', *mesa de centro* 'coffee table') (*Diccionario de americanismos*\(^{21}\)). Valencia and Echeverría’s (1999) Chilean homeland monolingual lexical availability data do, however, include both *coche*\(^{22}\) and *mesita*, although with low AIs, particularly for the latter (contrastively, *esquinero* 'end table' had an AI of 28 and *mesa de centro* 'coffee table' had an AI of 12). The item *ordenador* 'computer' was not found in the homeland data (but *computador* 'computer' was, showing conceptual awareness). Possible Peninsular dialectal influence was also observed in other categories, such as *culo* in *Human Body* (exclusively produced by the G2) and *patata* in *Food and Drink* (produced by both groups).

\(^{20}\) Two of the tokens were found in *Furniture*. To reflect an accurate reading of its extension in the corpus, another six tokens from *Science and Technology* were included in the count. The five G2 tokens were in variation with their other nine tokens of *computador* and six tokens of *computadora*. In addition to the G1’s three *ordenador* tokens, there were 2 of *computador* and 3 of *computadora*.

\(^{21}\) The *Diccionario de americanismos* contains no entry for *mesita* or any variation of it (e.g., *mesita de noche*), whereas the *Diccionario de la Real Academia Española* does.

\(^{22}\) Its low availability for homeland monolingual Chileans is evidenced by its inferior ranking to items like *triciclo* 'tricycle', *burro* 'donkey', and *cohete* 'rocket'.
Thus, the mere presence of some Peninsularisms as well as the rate with which others are mentioned in the group productions strongly suggest dialect contact effects. In addition, in tracing which speakers produced the 27 tokens, results revealed that all were produced by speakers who had traveled to Spain at least once. Twenty of the 27 tokens were produced by speakers who had traveled 2 or more times to Spain. The data suggest that regular, repeated short-term travel exposure to a non-heritage dialectal region can exert an influence\textsuperscript{23} on the lexicons of diaspora Spanish speakers, both of the first and second generation. Lastly, the prediction that the number of Peninsularisms would be proportionately smaller in the G1 data was confirmed, with 0.4 Peninsularisms per speaker versus 0.7 for the G2, although the difference was slight.

4.6 Research Question 2d: Conceptual Patterns

I now turn to the results of the analysis of cross-generational conceptual patterns in each of the six domains. Beginning with 


\begin{table}[h]
\centering
\begin{tabular}{|l|l|}
\hline
\textbf{Human Body} & \textbf{Table 4-19} displays the Top 15 items arranged according to AI. We can recall that the formula used to generate a lexical item’s AI takes into account both its frequency and order of mention across participants. Items with a higher AI are believed to be more prototypical members of the category in question.

\textsuperscript{23} As discussed in Chapter 2, media from Spain is frequently broadcast on Swedish public television, constituting a factor that may also lead to the adoption of Peninsularims. No data were collected for this factor, however.
Table 4-19
Top 15 Words in Comparison (Human Body)

<table>
<thead>
<tr>
<th>Rank of AI</th>
<th>G2</th>
<th>Difference</th>
<th>G1</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>mano</td>
<td>1</td>
<td>cabeza</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>corazón</td>
<td>7</td>
<td>mano</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>cabeza</td>
<td>-2</td>
<td>pie</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>pie</td>
<td>1</td>
<td>ojo</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>ojo</td>
<td>1</td>
<td>pierna</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>dedo</td>
<td>2</td>
<td>brazo</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>nariz</td>
<td>0</td>
<td>nariz</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>boca</td>
<td>11</td>
<td>dedo</td>
<td>-2</td>
</tr>
<tr>
<td>9</td>
<td>pelo</td>
<td>30</td>
<td>corazón</td>
<td>-7</td>
</tr>
<tr>
<td>10</td>
<td>pierna</td>
<td>5</td>
<td>oreja</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>estómago</td>
<td>9</td>
<td>riñón</td>
<td>53</td>
</tr>
<tr>
<td>12</td>
<td>brazo</td>
<td>-6</td>
<td>hígado</td>
<td>81</td>
</tr>
<tr>
<td>13</td>
<td>oreja</td>
<td>3</td>
<td>pulmón</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>uña</td>
<td>0</td>
<td>uña</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>cerebro</td>
<td>1</td>
<td>sangre</td>
<td>11</td>
</tr>
</tbody>
</table>

Compatibility - 80%

For the category of Human Body, there was an extremely high overall degree of convergence between the groups’ Top 15 available words (80%). Of the items that differed, three (boca 'mouth', pelo 'hair', estómago 'stomach') were exclusive to the G2, and another four were exclusive to the G1 (riñón 'kidney', hígado 'liver', pulmón 'lung', sangre 'blood'). These seven words, however, were not entirely absent in the opposite groups’ productions; they were simply further down in ranking due to their inferior AIs (occupying slots ranging from 19 [boca] to 92 [hígado]). The greatest ranking discrepancies between the two lists were tied to the G1 words hígado and riñón, which, along with sangre, suggest a stronger representation of internal elements in G1 conceptualizations of the body (at least in Spanish mode).

The preeminence of cabeza 'head' and mano 'hand' in both the G1 and G2 responses was clear, yet the high index of corazón 'heart' in the G2 data was distinctive. Interestingly, Valencia and Echeverría’s (1999) corresponding data for homeland speakers mirrors the G1 result: in their
data, corazón 'heart', ranked 11th, does not appear among the most immediate mentions for this category. Thus this word seems to have special prototypical status for the G2. Despite these few differences, the basic, material nature of Human Body results in highly similar conceptualizations. Figure 4-7 provides a visual depiction of the foregoing contrasts.

Figure 4-7. Relative prominence of top 15 ranked words (Human Body)

Table 4-20 presents the data corresponding to Furniture, which reveals a high coincidence in the primary items of this category.

<table>
<thead>
<tr>
<th>Rank</th>
<th>G2</th>
<th>Difference</th>
<th>G1</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>mesa</td>
<td>0</td>
<td>mesa</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>silla</td>
<td>0</td>
<td>silla</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>sillón</td>
<td>2</td>
<td>sofá</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>cama</td>
<td>0</td>
<td>cama</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>sofá</td>
<td>-2</td>
<td>sillón</td>
<td>-2</td>
</tr>
<tr>
<td>6</td>
<td>lámpara</td>
<td>-</td>
<td>velador</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>tavla</td>
<td>-</td>
<td>antiguo</td>
<td>20</td>
</tr>
</tbody>
</table>
With a slight difference in ranking order, the five most available items in the two lists are identical (mesa 'table', silla 'chair', sillón 'arm chair', cama 'bed', sofá 'couch'); 58% to 79% of the G2 participants produced these items while only 27% to 43% of the G1 did, which suggests that they carry more prototypical value for the younger group as a whole. Beyond these basic elements, the additional items vary drastically, to the extent that among the remaining words there is only one instance of (similarly ranked) overlap (comedor 'dining table/set'). Such difference results in an overall 20% between-group compatibility of items comprising the Top 15. Meriting mention is the G2’s greater emphasis on topically peripheral decorative or descriptive items, like lámpara 'lamp', espejo 'mirror', and reloj 'clock', whereas the G1 list contains more topically core items such as ropero 'closet', cómoda 'dresser' and vitrina 'glass cabinet'. The items velador 'night stand' and antiguo 'antique, old' were highly prototypical for the G1, but much less so for the G2. These differences are also represented in Figure 4-8.
As categories become more socially oriented and culturally imbued, the data reflect a greater degree of conceptual discontinuity. Table 4-21 displays the data for the domain of Modes of Transportation.

### Table 4-21
**TOP 15 WORDS IN COMPARISON (MODES OF TRANSPORTATION)**

<table>
<thead>
<tr>
<th>Rank</th>
<th>G2</th>
<th>Difference</th>
<th>G1</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>auto</td>
<td>0</td>
<td>auto</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>bus</td>
<td>0</td>
<td>bus</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>tren</td>
<td>0</td>
<td>tren</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>avión</td>
<td>1</td>
<td>bicicleta</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>coche</td>
<td>-</td>
<td>avión</td>
<td>-1</td>
</tr>
<tr>
<td>6</td>
<td>bicicleta</td>
<td>-2</td>
<td>barco</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>taxi</td>
<td>7</td>
<td>metro</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>autobús</td>
<td>-</td>
<td>bote</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>carro</td>
<td>-</td>
<td>micro</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>caminar</td>
<td>-</td>
<td>camión</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>moto</td>
<td>0</td>
<td>moto</td>
<td>0</td>
</tr>
</tbody>
</table>
There is a high degree of initial correspondence: four out first five items, and seven out of the Top 10, match. Yet there is evidence of conceptual distinctions. For example, there is an earlier and more frequent mention of taxi in the G2 list, possibly due to their more active lifestyles and urban upbringing. In addition, for the G1, items denoting water travel (barco 'ship/boat' and bote 'boat') were more prototypical transportation means. Another contrast observed related to non-machine means of travel; while the human means of caminar 'walk' and correr 'run' are mentioned by the G2, the G1 list includes animal travel (caballo 'horse'; burro 'donkey' and camello 'camel' make the Top 50). Figure 4-9 provides a side-by-side visual of the groups’ Top 15 configurations.

Figure 4-9. Relative prominence of top 15 ranked words (Modes of Transportation)
Turning to the domain of *Professions and Trades*, we observe a great deal of variability even in the uppermost slots (Table 4-20).

<table>
<thead>
<tr>
<th>Rank</th>
<th>G2</th>
<th>Difference</th>
<th>G1</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>profesor</td>
<td>0</td>
<td>profesor</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>enfermera†</td>
<td>6</td>
<td>carpintero</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>policia</td>
<td>-</td>
<td>chofer</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>doctor</td>
<td>0</td>
<td>doctor</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>bombero</td>
<td>-</td>
<td>médico</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>mecánico</td>
<td>5</td>
<td>arquitecto</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>vendedor</td>
<td>20</td>
<td>ingeniero</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>abogado</td>
<td>2</td>
<td>enfermera†</td>
<td>-6</td>
</tr>
<tr>
<td>9</td>
<td>médico</td>
<td>-4</td>
<td>cocinero</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>juez</td>
<td>-</td>
<td>abogado</td>
<td>-2</td>
</tr>
<tr>
<td>11</td>
<td>chef</td>
<td>-</td>
<td>mecánico</td>
<td>-5</td>
</tr>
<tr>
<td>12</td>
<td>bibliotecaria†</td>
<td>-</td>
<td>maestro</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>conductor</td>
<td>30</td>
<td>secretaria†</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>piloto</td>
<td>-</td>
<td>modista</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>programador</td>
<td>-</td>
<td>jugador de fútbol</td>
<td>-</td>
</tr>
</tbody>
</table>

Compatibility - 40%

†Items bear feminine morphological marker because this was the sole form in which they were produced.
For both groups profesor 'professor' was clearly the most prototypical profession, and enfermera 'nurse' and mécanico 'mechanic' also coincided, though not in ranking. A significant conceptual contrast between the two lists was the presence of common public service occupations (policía 'police officer', bombero 'fireman', juez 'judge') in the G2 list and their complete absence in the G1 list. Given their high visibility and familiarity in society, such an absence in the G1 list is quite remarkable; it is seemingly tied to the exile experience of many and their common characterization as averse to public authority. Furthermore, blue-collar manual jobs were notably more prominent in the G1 list (carpintero 'carpenter', chofer 'chauffeur, driver', cocinero 'cook', modista 'tailor'), suggesting divergent experiences and thus conceptualizations of employment between them and their offspring. Relatedly, an interesting contrast is the G2's more prestigious chef 'chef' compared to the G1's cocinero 'cook'. These contrasts are depicted in Figure 4-10.

Figure 4-10. Relative prominence of top 15 ranked words (Professions and Trades)
For the domain of Economy, Table 4-23 shows a 50% overlap in the groups' Top 10 responses and 43% overall compatibility, with the highly prototypical dinero 'money' occupying the primary slot cross-generationally. The concept of rico 'rich' ranked higher for the G2 than that of pobre/pobreza 'poor/poverty'. Likewise, the concepts of trabajo 'work' and ahorros 'savings' were more prototypical for the G1, perhaps reflective of distinct perspectives or life stages. The G2's production of the general classificatory term moneda (whether meaning 'coin' or 'currency') can be contrasted with the G1's homeland focused peso 'peso'. These contrasts are depicted in Figure 4-11.

TABLE 4-23
TOP 15 WORDS IN COMPARISON (ECONOMY)

<table>
<thead>
<tr>
<th>Rank</th>
<th>G2</th>
<th>Difference</th>
<th>G1</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>dinero</td>
<td>0</td>
<td>dinero</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>rico†</td>
<td>-</td>
<td>pobreza†</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>plata</td>
<td>-</td>
<td>trabajo</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>malo</td>
<td>0</td>
<td>malo</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>pobre†</td>
<td>-</td>
<td>difícil</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>moneda</td>
<td>-</td>
<td>cuenta</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>bueno</td>
<td>2</td>
<td>ahorros</td>
<td>44</td>
</tr>
<tr>
<td>8</td>
<td>problema</td>
<td>-</td>
<td>riqueza†</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>finanzas</td>
<td>-</td>
<td>bueno</td>
<td>-2</td>
</tr>
<tr>
<td>10</td>
<td>trabajo</td>
<td>-7</td>
<td>sueldo</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>estudios</td>
<td>-</td>
<td>desarrollo</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>banco</td>
<td>1</td>
<td>peso</td>
<td>33</td>
</tr>
<tr>
<td>13</td>
<td>planificar</td>
<td>-</td>
<td>banco</td>
<td>-1</td>
</tr>
<tr>
<td>14</td>
<td>impuestos</td>
<td>0</td>
<td>impuestos</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>pobreza†</td>
<td>-13</td>
<td>recorte</td>
<td>-</td>
</tr>
</tbody>
</table>

Compatibility - 43%
Finally, the domain of Politics (Table 4-24) presents a fairly high conceptual convergence between groups (47%), with presidente 'president' as the most prototypical item cross-generationally, terms related to the political continuum (partido 'party', derecha 'right', izquierda 'left'), and terms denoting political office (ministro 'minister', gobernador 'governor', diputado 'representative'). It is interesting that, for both groups, the international presidente 'president' ranks higher than ministro 'minister', despite the latter being the office of Sweden's leader. Aside from neutral, classificatory terms, both lists also contain a significant number of negatively connoting items: at least three in the G2 list (problema 'problem', mentiroso 'liar', mentira 'lie') and four in the G1 list (mentira 'lie', corrupción 'corruption', sucia 'dirty', ladrón 'thief') evoke issues of abuse and distrust, though those corresponding to the G1 are especially hostile and of high prototypicality.
TABLE 4-24
TOP 15 WORDS IN COMPARISON (*POLITICS*)

<table>
<thead>
<tr>
<th>Rank</th>
<th>G2</th>
<th>Difference</th>
<th>G1</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>presidente</td>
<td>0</td>
<td>presidente</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>dinero</td>
<td>-</td>
<td>mentira</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>problema</td>
<td>-</td>
<td>partido</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>gobierno</td>
<td>30</td>
<td>senador</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>país</td>
<td>-</td>
<td>corrupción</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>derecha</td>
<td>6</td>
<td>ley</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>partido</td>
<td>-4</td>
<td>sucia</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>elección</td>
<td>49</td>
<td>ministro</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>izquierda</td>
<td>23</td>
<td>promesa</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>ministro</td>
<td>-2</td>
<td>político</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>mentiroso</td>
<td>-</td>
<td>diputado</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>democracia</td>
<td>-</td>
<td>derecha</td>
<td>-6</td>
</tr>
<tr>
<td>13</td>
<td>derecho</td>
<td>-</td>
<td>ladrón</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>mentira</td>
<td>-12</td>
<td>demócrata</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>ley</td>
<td>-9</td>
<td>gobernador</td>
<td>-</td>
</tr>
</tbody>
</table>

Compatibility - 47%

While it is clear that the G2 as a group share to a large extent the G1's conception of politics, the data suggest a more negative outlook on the part of the G1. This result seems in line with the especially painful and turbulent politically motivated experiences of many of the G1 participants relative to their Swedish-born or raised children. A look at Valencia and Echeverría's (1999) homeland data for this domain reveals a clear contrast: among the Top 15 ranked items, not one is negatively connoting. Thus the present data evidence a diaspora-specific conception of this theme. A visual depiction of these differences is found below in Figure 4-12.
Figure 4-12. Relative prominence of top 15 ranked words (Politics)

The foregoing data in this section provide a sampling of cross-generational conceptualizations in a minority language setting. Contrasting life experiences and patterns of socialization appear to result in certain subtle yet identifiable organizational and prototypical discontinuities in the mental lexicon. The Top 15 words with the highest AIs were studied as indicators of each group’s domain conceptualizations. Taking into account the percentage of between-group compatible words within these uppermost slots, the pattern that emerged was the following:

80% Human Body  
67% Modes of Transportation  
47% Politics  
43% The Economy  
40% Furniture  
40% Professions and Trades

The groups’ conceptualizations were highly similar in both Human Body and Modes of Transportation, domains belonging to the Basic and Community supracategories. Surprisingly, the more socio-culturally laden Politics generated the next highest compatibility rate, above even
the more basic *Furniture* and *Professions and Trades*; in this chapter, not only have we seen that
the G2 produced as many words as the G1 in *Politics* but now also that their conception of the
topic is quite similar. Such conceptual continuity appears to align with Becker (2013) and
Gibbons and Ramirez’s (2004), who found that exile-background SHS's frequent engagement in
political discourse in family circles provided important opportunities for linguistic and cultural
transmission. That is, frequent explicit discussion as well as many G2s' strong identification with
an exile heritage were likely contributing factors to the striking degree of cross-generational
conceptual convergence in this domain, despite the G2's socialization in the Swedish political
environment.
5. CONCLUSIONS

5.1 Primary Research Conclusions

This project investigated quantitative and qualitative aspects of the lexical repertoires of two generations of Spanish-Swedish adult bi-/trilinguals residing in Stockholm, Sweden, including issues related to HL development and language change as a result of cultural and linguistic contact. In addition, certain aspects of the speakers' lexical productions were also contrasted with comparable homeland and U.S. HS data.

With respect to quantitative variation, the study found that in only a small minority of domains (three out of 21) did the between-group word production differ significantly. That is, the average number of words produced overall was remarkably similar between the groups. Although the G2 advantage was noted in what was classified as a basic domain (Spirituality and Religion) and the G1 advantage in less domestic domains (Professions and Trades and Countryside), when performance was compared more broadly across the three supra-categories (Basic, Community and Society), no statistical difference was found. Thus, domain complexity did not significantly predict between-group differences in the quantitative aspects of word production. This result is in line with Victery's (1971) lexical availability study, which reported equal or slightly higher word totals for bilinguals in Texas than for their monolingual Spanish counterparts in both Texas and Mexico. These results challenge the common claim that HS lexicon is largely limited to domestic domains.

A comparative analysis of the quantitative data with corresponding homeland means for 14 domains revealed that both the G1 and G2 means were significantly lower. In fact, in only two domains (Furniture and The Environment) did the G1 means exceed those of the homeland group, and in only one (Furniture) did the G2 slightly do so. The diaspora speakers' general advantage over homeland speakers in Furniture suggests that its development is favored in the context of a Swedish
cultural backdrop. Sweden’s advanced preoccupation with and stakehold in the houseware and home furnishing industry may be responsible for increased conversations about and exposure to this conceptual domain for minority language speakers. Comparative qualitative analyses of homeland versus diaspora speaker productions in this domain could shed more light on the observed quantitative differences. The general diaspora-homeland divide (which strongly favors homeland speakers) supports previous research showing age-divergent linguistic trajectories for HS (homeland data was collected among high school aged students), but also points to possible lexical attrition on the part of the G1. Lexical lacunae, lexical loss, and problems with lexical access may all be responsible in varying degrees for such incongruities in word production. While lexical availability research is able to shed some light on HS lexical lacunae versus slowed lexical access (via the complete absence versus mere late mention of items from group to group), future research should work to further tease apart these different phenomena affecting the lexicons of diaspora speakers.

An additional finding was that response totals were notably more variable among G2 participants than among their G1 counterparts, suggesting that, in addition to widely varying competencies in other grammatical areas, HS are characterized by a broad proficiency range when it comes to the lexicon as well. Similarly, a comparison of lexical diversity between the two participant groups and SHS in Chicago found that the word lists of both HS groups were more diverse (i.e., there was less agreement among participant responses) than the G1 word lists, signaling highly variable networks in the mental lexicon. According to Schmitt (2010), a lack of systematicity across speaker responses on an association task is a sign of immature, weakly organized lexical networks. It appears, then, that a less obvious effect of reduced input on the lexicon has to do with depth of word knowledge, which includes semantic and other associations
among words. Thus, there is a clear need for word families and word relationships to take center stage in the vocabulary instruction of SHS courses.

With respect to the G2 social variables tested, two predictions were confirmed while another two were not. First, exile background emerged as the most discriminating variable; HS whose parents migrated due to political persecution demonstrated stronger knowledge in 15 of the 24 possible domains and categories. Second, for the variable of gender, the males produced significantly more words in domains pertaining to the Community and Society supra-categories, whereas in Basic the genders performed similarly. This result contrasts with lexical availability work in majority language contexts that generally shows a female advantage, but aligns with the study predictions based on research in minority language contexts, particularly in the Chilean-Swedish context, suggesting greater heritage ethnolinguistic affiliation on the part of male bilingual youth. Third, parental education was not significantly correlated with word production. This factor appears to behave differently than in studies of lexical availability in majority Spanish contexts, where parental education and, relatedly, SES are often strong predictors of children's vocabulary development. This finding represents a contribution to the field's limited and inconclusive knowledge of the role of SES in HL maintenance and development. Finally, the results for the variable of birth order did not fully align with predictions based on previous research in minority language contexts showing that the higher the birth order, the stronger the proficiency. The small n's for each birth order group may explain the lack of significance.

With respect to the relationship between the various language variables and total word production, HL education and Spanish proficiency were found to be significant (or nearly significant) predictors for the G2, although self-reported Spanish proficiency, use, and media consumption were not. These results showed that the external proficiency measure and non-evaluative data (amount of instruction) were better predictors of lexical knowledge than evaluative
data like self-reports. Importantly, these results support the idea that lexical tests may reliably predict global language proficiency for SHS, though the correlation was lower than what Fairclough (2011, 2013a) has reported in her work based on different measures. They also show that HL instruction, even at approximately an hour a week, can result in significant linguistic benefits, although possibly only short-term, as Bylund and Diaz (2012) found in this same community. In addition, both English proficiency and self-reported English proficiency were positively correlated with total words, signaling the L3’s positive effect on HL development. For the G1, Spanish proficiency also predicted total words, as well as self-reported Spanish proficiency (suggesting that G1 speakers are more accurate judges of their abilities).

In addition, an analysis of the compatibility between the groups' lists in six domains revealed an average overlap of 20%. Basic, closed categories (such as Human Body) exhibited the highest degree of overlap. The lowest rate of overlap surprisingly corresponded to Professions and Trades, a seemingly straightforward category. Not only did the groups differ significantly in the number of words offered in this category (the G1 produced significantly more), as seen for RQ1, but also in the actual words supplied. In fact, a Root TTR analysis of all six domains showed once again that the G2 consistently produced a wider variety of words than the G1, showing less agreement amongst themselves in lexical/conceptual norms and association patterns. Thus, we can again conclude that there is a strong need for HS instruction to incorporate vocabulary activities that address lexical relations, including etymological, semantic (e.g., paradigmatic, metonymous, synonymous, etc.), and stylistic connections. Given that the study identified Professions and Trades as a primary field in which HS knowledge diverges sharply with that of more proficient speakers, it is one that merits special focus in bilingual education for younger heritage learners, in SHS courses, and perhaps also in Business Spanish courses with high HS enrollment. The Countryside appears to be another domain in which HS know few words relative to their parents.
In studying the types of words produced cross-generationally, four main sources of difference were observed. First, there were stylistic differences with the G2 more often producing informal, colloquial words (guata 'belly'; fome 'stupid, boring') and the G1 more likely to provide technical, higher register words (cráneo 'cranium'; diputado 'representative'). Relatedly, there were word frequency differences with lower frequency words more abundant in the G1 data (fuera 'strength') than in the G2 data (fuerte 'strong'). These observations support the widely accepted notion that HL instruction should focus on broadening speakers' stylistic range and familiarity with a variety of literary genres. Third, the G2 productions contained a greater proportion of syntagmatic (different word class), as opposed to paradigmatic (same word class), responses (e.g., complicado 'complicated', horrible 'horrible' in POL). According to Schmitt (2010), a larger number of syntagmatic responses, which rely on syntactic rather than semantic word relationships, can signal less mature semantic networks in the mental lexicon. Lastly, among the G2 responses was a proportionately higher number of semantically peripheral items (reloj 'clock', almohada 'pillow' in Furniture), possibly a result of lexical lacunae. Despite these patterns of difference, this portion of the analysis confirmed the G2's robust lexical store. Not only did a significant portion of the G2 words coincide with G1 productions, the G2 also independently produced many complex, technical, and semantically rich words.

The characteristics of the evolving variety of the Chilean-Swedish community were also analyzed by way of studying the rate and types of cross-linguistic transfer in the data. For both groups, Swedish borrowings were more abundant than English borrowings, an order mirroring the relative societal and cognitive importance of each language. However, although the G2 was responsible for a greater percentage of both kinds, borrowings represented only a tiny fraction of the total lexical items (between 0.7% and 1.5%). For both participant groups and both borrowing types (Swedish and English), the prediction that borrowings would be more prevalent in the advanced
domains was confirmed. The prediction that the G2 would produce more borrowings than the G1 in the basic domains was also confirmed, showing their more intense language contact. The minimal presence (or total absence, in the case of Human Body) of borrowings in the basic domains, however, confirmed the strong Spanish proficiency levels of the participants, since extensive borrowing at the basic level may denote an advanced stage of language shift (Hock & Joseph 2002; Nagy 2011). Interestingly, the domain for which most Swedish and English borrowings were produced cross-generationally was Professions and Trades, seemingly linked to the G2’s lexical gaps (e.g., chemist) but also to the G1’s wrestle with novel (post-exile) professions (e.g., computer technician).

Regarding the different borrowing phenomena, loanwords and loanblends were the most plentiful in the data, in line with previous language contact research. While the G2 was responsible for more phonological intrusions, likely tied to their low Spanish literacy, the G1 surprisingly produced a greater proportion of calques, which tend to be more frequent among G2 speakers. Additionally, while for the G2 established borrowings were more plentiful than lone/nonce borrowings, the opposite was true for the G1. Overall, the impression from the foregoing borrowing data is that Swedish and English contact are responsible for very little change in the lexicons of the participants.

The study also measured the presence of lexical items specific to Peninsular Spanish varieties. Three types (27 tokens) were noted in three of the six domains evaluated. Although numerically few, their availability (i.e., token frequency and order of mention) across speakers were striking. The variants were produced exclusively by those participants who had spent time in Spain, especially those who had made recurrent visits. This was taken as evidence of dialect contact effects resulting from the participants’ repeated short-term travel to this dialectal zone. The prediction that the G2 data would contain a higher rate of Peninsularisms was confirmed, although the difference
was very small. This is significant because it suggests that age constraints generally influencing dialect contact may also apply to HS whose dominant language is generally not the HL.

Lastly, with respect to conceptual patterns, the data suggest that language maintenance and lexical knowledge do not necessarily imply conceptual continuity. Despite similar cross-generational word production rates, it was shown for example, that the G1's conceptualization of *Modes of Transportation* included animals, and that for the G2 *taxi* was a more prototypical item. Likewise, in *Professions and Trades* we observed the high prototypically of manual labor jobs and low prototypicality of public service jobs for the G1 relative to the G2. In *Economy*, the concept of wealth had higher prototype status for the G2. In *Politics*, we noted the greater prominence of negatively connoting lexical items for the G1. It would seem, then, that cross-generational conceptual diversity in the mental lexicon can, in some respects, be viewed as existing in a separate dimension from that of language shift. Language contact research has long shown that cultural contact resulting in conceptually novel lexical phenomena is not necessarily synonymous with language loss (Otheguy & García 1993; Otheguy & Stern 2011). Likewise, the present analysis explores how variable world perceptions and culturally induced conceptualizations can shape the categorical structure of the mental lexicon, a previously unexamined aspect of lexical variation in the Spanish-speaking diaspora.

### 5.2 Study Significance

This project contributes to theoretical understandings of HL acquisition through an examination of lexical knowledge, an area of the linguistic systems of HS that has received relatively little attention. Specifically, this study nuances the field's understanding of the quantitative and qualitative nature of the productive lexical knowledge of HS relative to that of Spanish dominant and monolingual speakers both in their diaspora community and in Chile, their family's country of origin. By focusing on thematically based vocabulary knowledge, insights have been gained into the
kinds of input HS have received within their families and communities, as well as the types of linguistic tasks they may be able to realize. Previous depictions of HS lexicon have characterized it as strongly restricted to the domestic realm, yet adequate explanations and data to support this claim have been missing.

This study has tested various aspects of HS lexicon in identifiable thematic domains existing along the domestic to extradomestic continuum. In addition, to further understand the nature of the HS participants' lexical development, the relationship between lexical availability and several community relevant social and linguistic factors was examined. Not only has this project focused on an understudied Spanish-speaking community, the impact of the acquisition of a societally prestigious L3 (English) on SHL maintenance has for the first time been addressed. Given the data presented in this study that confirms the thematically diverse lexical knowledge of HS in this community, an important question to ask is whether it is appropriate to classify, regard, or label SHL solely as a "home" language. Further, we might ask what the classification "home" is meant to signify, and whether there is a term that can more aptly capture the extradomestic, and even extracommunity, reach of the HL for HS. Perhaps simply referring to the HL as a "family" or "community" language is sufficient.

This project has also contributed novel perspectives to the borrowing patterns of minority language speakers in the context of two high status societal languages (English and Swedish). Also, along the lines of language change and variation, this study has used a Cultural Linguistics framework to show how linguistic analysis can provide a window into cross-generational conceptual change in a diaspora community. A more covert area of linguistic variation, certain group-specific conceptual patterns in the mental lexicon may exist largely outside of proficiency concerns. It is hoped that the findings presented on this topic will spur further research into culturally induced conceptual discontinuities across different speaker groups.
Naturally, this project also has pedagogical importance. According to Valdés (2001), “for many educators engaged in the teaching of Spanish to heritage Spanish speakers, Spanish language maintenance is an important and primary goal” (p. 51). She has established that among the goals of a well-structured HS language track is the acquisition of a prestige variety of Spanish (1997), and although there is no common consensus on which variety that should be (Villa 1996), lexicon is undoubtedly a main factor. Likewise, Brecht and Ingold (2002, p. 2) note that a principal objective of the Center for Applied Linguistics Heritage Language Initiative is to build an education system “capable of producing a broad cadre of citizens able to function professionally in both English and another language.” Angelelli and Kagan (2002) argue that Superior-level attainment (according to the ACTFL scale) of HS with advanced abilities is a “realistic expectation that can be achieved through institutionally-based direct instruction” (p. 199). One prerequisite for the success of such classroom endeavors is that the linguistic systems of HS be well understood to provide a clear articulation of their educational needs. Such information can support justification for and guidance in the curricular development of SHS programs. Vocabulary is a fundamental and increasingly prioritized aspect of language instruction that is especially key in promoting the register expansion of HS. In fact, Kagan (2009, p. 21) has estimated that “increasing vocabulary may emerge as the primary goal of HLL instruction.” Efforts must be made for lines of communication to be established and remain open between researchers and practioners to maximize the integration of research findings into course curricula. If research can help identify particularly strong thematic areas and associated skills in HS lexical inventories, educators can work to strategically bridge these with other less familiar domains.

5.3 Study Limitations

There were several limitations to this project. Concerning the methods employed, one had to do with the written format of the association task. This medium may have served to intimidate the HS
who, according to the survey data, rate themselves quite low in their writing and reading skills in Spanish. It may also have been taxing for the older generation, who may not have been accustomed to or who may not have felt fully physically equipped to write so much within such a short amount of time. Cremer et al. (2011) submit that a spoken association task "is likely to tap into relations between words in the mind in a more direct way" (p. 190). However, other research has shown that lexical availability tasks administered orally tend to generate more colloquial and tangential vocabulary. In addition, while lexical availability data are produced in response to contextualized prompts and are meant to represent the vocabulary that speakers could recall under a given set of circumstances, the process of production is still far from naturalistic. There are many other factors (environment, interlocutor, familiarity, etc.) that can influence a speaker's word choices in everyday language use. Relatedly, the lexical availability task favors the production of nouns (because of their semantic weight), such that the data collected tends to be quite noun-heavy. Thus, it may be more accurate to talk about nouns, rather than the more inclusive lexicon.

Also in relation to this association task, it would have benefited the participants to complete an additional round or two of practice with subsequent feedback. Although the guidelines were discussed and repeated on each sheet, some participants still violated certain ones, such as producing proper nouns (that were later omitted in the analysis phase). Finally, had the participant sample sizes been larger, stronger lexical patterns would likely have emerged and perhaps more certitude attributed to those that did. As discussed in the introduction, a primary methodological challenge in studying lexical variation is the difficulty of obtaining a sufficient number of tokens of the items in variation, which can be partly mitigated through large participant samples.

Another issue was the use of the DELE as an external proficiency measure of Spanish. I believe it would have been more suitable to use a test that (1) employed a general or Chilean variety of Spanish, rather than Peninsular Spanish; and (2) tested more specific grammatical concepts.
Another main drawback of the study was the relatively small sample sizes. Given the minority status of Chileans in Sweden and their now scattered residential patterns, participant recruitment posed difficulties. Also, the fact that the samples size differed so extremely from that of the homeland data affected the reliability of those specific comparisons.

5.4 Future Directions

Regarding the lexicon of HS, research is needed that studies it from a variety of angles. For example, the relationship between lexical proficiency and global language proficiency should be explored in more detail, especially due to its ability to inform best practices in instruction and to streamline diagnostic or placement procedures. It would be of interest to examine which aspects of grammatical knowledge are most strongly correlated with different measures of lexical knowledge.

As Fairclough (forthcoming) points out, "since vocabulary knowledge appears to be explicit for all learners (including HS), L2 vocabulary research, methods and materials can be transferrable to the [heritage learner] context" (p. 19). She suggests that future study of the lexicon of SHS should be guided by L2 research with a focus on topics such as "multi-word lexical items, vocabulary richness, comparison of the lexicon in different types of registers, depth of knowledge (e.g., word associations)," among others. Research should examine the effectiveness of different types or combinations of vocabulary learning strategies. Schmitt (2010, pp. 26-28) provides a list of nine factors that facilitate vocabulary acquisition, or engagement with vocabulary. He argues that the following aspects should be (more fully) addressed in instructional contexts:

(1) frequency of exposure
(2) attention focused on lexical item
(3) noticing of lexical item
(4) intention to learn lexical item
(5) requirement to learn lexical item (by teacher, test, syllabus)
(6) a need to learn/use lexical item (for task or for a personal goal)

(7) manipulation of lexical item and its properties

(8) amount of time spent engaging with lexical item

(9) amount of interaction spent on lexical item.

Outside of the classroom, naturalistic or interview data should be utilized or sought so that features of HS lexicon can be observed in a range of highly situated, dialogic conditions. The SPINTX (http://spanishintexas.org) and CHISPA (Potowski & Torres, forthcoming) corpora are two examples of resources containing HS speech obtained through sociolinguistic interviews. Another area to look to for insightful data on HS lexical use and development is that of language brokering in healthcare, legal and other contexts. HS of many ages routinely serve as informal interpreters for family members and in so doing acquire sophisticated skills, such as breaking words into parts, that influence vocabulary development (Belpoliti & Plascencia-Vela 2013).

With specific regard to the SHS of the Chilean-Swedish community, I will mention a few research directions. While this study looked at the independent influence of English and Swedish on the HL lexicon, given their etymological similarities it would be important to examine the ways in which the two languages interact to jointly influence the Spanish spoken in this community. This study has also identified Peninsular Spanish lexical items among the productions of both the G1 and G2 participants; to complement this finding, lexical recognition tasks or attitudinal experiments could be designed to further examine these speakers familiarity with and perceptions of Peninsular vocabulary. Similarly, perceptual studies would be useful in determining the G2's register sensitivities with respect to the lexicon. For example, it would interesting to explore the level of formality speakers attribute to colloquial items like guata 'belly' or plata 'cash' versus other known variants like estómago 'stomach' or dinero 'money', as well as what the relationship between their perceptions and patterns of use looks like.
REFERENCES


Silva-Corvalán, C. 2003. Linguistic consequences of reduced input in bilingual first language acquisition. In S. Montrul and F. Ordóñez (Eds.). Linguistic theory and


Publishers.


APPENDICES

APPENDIX A: Questionnaire

1. Cuestionario básico

Datos básicos
Sexo  M ____ F____
Edad __________
Posición en la familia (primer hijo, segundo, tercero, etc.) ____________________________
¿Dónde nació? (barrio, ciudad, región, país) ____________________________
¿Si nació en otro país, a qué edad llegó a Suecia? ____________________________
¿Dónde nacieron y crecieron sus padres (barrio, ciudad, región, país)?
Madre: ____________________________
Padre: ____________________________
¿Qué edad tenían sus padres al llegar a Suecia? ____________________________
¿Cuál fue la razón principal por la que Ud. o sus padres vinieron a Suecia?________________________
¿En qué barrios de Estocolmo ha vivido? (el barrio en que actualmente vive debe aparecer último)
¿Ha vivido en otra parte de Suecia o en otro país que no sea Chile? –haga una lista de todos los lugares:
¿Vive la mayoría de su familia en Suecia? ¿En qué otros lugares tiene familia?
¿Qué tan probable es que Ud. o miembros de su familia nuclear regresen a Chile a vivir? (1= poco probable, 5=casi seguro) 
1 2 3 4 5

Escuela
¿Dónde asistió a la escuela secundaria (gymnasieskolan) (nombre de la escuela, comuna, ciudad, país)?
¿Era una escuela pública o privada? ____________________________
Si asistió a la escuela en Suecia, ¿Participó en hemspråk? _______ ¿Por cuántos años? _______

Estudios y ocupaciones
¿Cuál es el ingreso mensual (aproximado) de su familia? ________________
¿Qué nivel de estudios tiene Ud.? (Circle one)

<table>
<thead>
<tr>
<th>Escuela primaria</th>
<th>Escuela media</th>
<th>Escuela secundaria</th>
<th>Título universitario</th>
<th>Título avanzado</th>
</tr>
</thead>
</table>

¿Cuál es su ocupación? ____________________________
¿Qué nivel de estudios tiene su madre?

<table>
<thead>
<tr>
<th>Escuela primaria</th>
<th>Escuela media</th>
<th>Escuela secundaria</th>
<th>Título universitario</th>
<th>Título avanzado</th>
</tr>
</thead>
</table>

¿Cuál es (ha sido) la ocupación principal de su madre? (Sea específico) ____________________________
¿Qué nivel de estudios tiene su padre?

<table>
<thead>
<tr>
<th>Escuela primaria</th>
<th>Escuela media</th>
<th>Escuela secundaria</th>
<th>Título universitario</th>
<th>Título avanzado</th>
</tr>
</thead>
</table>

¿Cuál es (ha sido) la ocupación principal de su padre? (Sea específico) ____________________________
¿Cómo describiría su nivel socioeconómico en Chile (o el de sus padres) antes de venir a Suecia? ¿Era superior, inferior, o igual al que tienen ahora? ____________________________

Lenguas
¿A qué edad empezó Ud. a aprender y a usar el sueco? ____________________________
¿A qué edad empezó Ud. a aprender y a usar el español? ____________________________
Primera(s) lengua(s) aprendidas de sus hermanos mayores: ____________________________
¿Sus hermanos mayores hablan mejor el español o el sueco? ___________________

Primer(a)s lengua(s) aprendidas de sus hermanos menores: ___________________

¿Sus hermanos menores hablan mejor el español o el sueco? ______________________

Entre usted y tus hermanos, ¿quién habla el mejor español? (El mayor, el/uno de en medio, el menor) __________________________________

Entre usted y tus hermanos, ¿quién usa el español más? ______________________

Español

¿Cómo calificaría su nivel de español con respecto a la lectura? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de español con respecto a la escritura? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de español con respecto a la comprensión? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de español con respecto a su habilidad de hablar? (1 = pobre, 5 = muy bien)
1 2 3 4 5

Indique la frecuencia actual con la que está expuesto al español en el contexto de leer. (1 = casi nunca, 5 = frecuentemente)
1 2 3 4 5

Indique la frecuencia actual con la que está expuesto al español en el contexto de mirar la televisión. (1 = casi nunca, 5 = frecuentemente)
1 2 3 4 5

Indique la frecuencia actual con la que está expuesto al español en el contexto de escuchar música. (1 = casi nunca, 5 = frecuentemente)
1 2 3 4 5

Inglés

¿Cómo calificaría su nivel de inglés con respecto a la lectura? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de inglés con respecto a la escritura? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de inglés con respecto a la comprensión? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de inglés con respecto a su habilidad de hablar? (1 = pobre, 5 = muy bien)
1 2 3 4 5

Indique la frecuencia actual con la que está expuesto al inglés en el contexto de leer. (1 = casi nunca, 5 = frecuentemente)
1 2 3 4 5

Indique la frecuencia actual con la que está expuesto al inglés en el contexto de mirar la televisión. (1 = casi nunca, 5 = frecuentemente)
1 2 3 4 5

Indique la frecuencia actual con la que está expuesto al inglés en el contexto de escuchar música. (1 = casi nunca, 5 = frecuentemente)
1 2 3 4 5

Sueco

¿Cómo calificaría su nivel de sueco con respecto a la lectura? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de sueco con respecto a la escritura? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de sueco con respecto a la comprensión? (1 = pobre, 5 = muy bien)
1 2 3 4 5

¿Cómo calificaría su nivel de sueco con respecto a su habilidad de hablar? (1 = pobre, 5 = muy bien)
1 2 3 4 5

Desde que vive en Suecia, ¿ha viajado a un país donde se habla español (incluido Chile)?

| Viaje | País visitado | ¿Por qué? (trabajo, vacaciones, | ¿Por cuánto tiempo? |
¿Está Ud. casado(a) o vive con una pareja? ¿Cuál es su etnicidad? ______________ ¿Habla español? ______

¿Cuál es la proporción de sus amigos más cercanos que hablan español? ______% 

¿Pertenece Ud. a alguna(s) asociación(es) o club(s) de barrio (o de otro tipo, ej. iglesia) en el que se usa español? ______________

¿Cuáles son las actividades típicas de la asociación? ____________________________________________

¿Son la mayoría de los miembros de la asociación chilenos? Si no, ¿de dónde son? ______________

2. Uso de la lengua

Responde por favor a las siguientes proposiciones con hacer un círculo en el número que refleja más apropiadamente su manera de usar la lengua.

5=siempre/casi siempre
4=usualmente
3=casi 50% del tiempo
2=raras veces
1=casi nunca o nunca
na=no se aplica

1- Hablo español con mis amigos hispanos de mi edad.

   1 2 3 4 5 na

2- Uso el español cuando uso el “internet”.

   1 2 3 4 5 na

3- Hablo español con mi padre en casa o en privado.

   1 2 3 4 5 na

4-Cuando voy al cine, o arriendo videos, escojo las películas de lengua española.

   1 2 3 4 5 na

5- En la escuela secundaria (högstadiet och gymnasieskolan), hablaba español con mis maestros.

   1 2 3 4 5 na

6- Hablo español con mis compañeros de trabajo.

   1 2 3 4 5 na

7- Hablo español con mi jefe / supervisor en el trabajo.
8- Hablo español con mi hermano(s) mayor(es) en casa o en privado.

9- Escucho la radio en español.

10- Miro programas de televisión, además de las noticias, en español.

11- Hablo español con mi hermano(s) menor(es) en casa o en privado.

12- Hablo en español con mi madre en lugares públicos (por ejemplo, en el supermercado, en el centro comercial, en los restaurantes, en el tren/metro, etc.)

13- Paso mi tiempo libre con otros hispanohablantes (sean bilingües o monolingües en español).

14- Miro las noticias en la televisión en español.

15- Hablo español con mi abuelo.

16- Hablo español con mi abuela.

17- Hablo español con mis hermano(s) menor(es) en lugares públicos.

18- Hablo español con los clientes donde trabajo.

19- Hablo español con mi padre en lugares públicos.

20- Voy a la iglesia donde se da el sermón en español.

21- Discuto la política en español.

22- Me hablo (en silencio) o sueño despierto en español.

23- Cuando duermo, sueño en español.

24- Cuento en español.

25- Cuando estoy solo, pienso en voz alta o me hablo en español.
26- Oro (rezos) en español.

1 2 3 4 5 na

27- Hablo español con mi madre en casa o en privado.

1 2 3 4 5 na

28- Leo los periódicos y las revistas en español.

1 2 3 4 5 na

29- Hablo español con mis hermano(s) mayor(es) en lugares públicos.

1 2 3 4 5 na

30- Hago una mezcla del sueco y del español en conversaciones informales con amigos, hermanos y/o familia.

1 2 3 4 5 na
APPENDIX B: Diplomas of Spanish as a Foreign Language (DELE)

Name ________________

Multiple Choice Test

Each of the following sentences contains a blank indicating that a word or phrase has been omitted. Select the choice that best completes the sentence.

1. Al oír del accidente de su buen amigo, Paco se puso ________.
   a. alegre        b. fatigado       c. hambriento       d. desconsolado

2. No puedo comprarlo porque me ____________ dinero.
   a. falta        b. dan           c. presta          d. regalan

3. Tuvo que guardar cama por estar ____________.
   a. enfermo       b. vestido       c. ocupado         d. parado

4. Aquí está tu café, Juanito. No te quemes, que está muy ____________.
   a. dulce        b. amargo       c. agrio          d. caliente

5. Al romper los anteojos, Juan se asustó porque no podía ____________ sin ellos.
   a. discurrir     b. oír           c. ver             d. entender

6. ¡Pobrecita! Está resfriada y no puede ____________.
   a. salir de casa  b. recibir cartas  c. respirar con pena  d. leer las noticias

7. Era una noche oscura sin ________.
   a. estrellas     b. camas         c. lágrimas        d. nubes

8. Cuando don Carlos salió de su casa, saludó a un amigo suyo: -Buenos días, ____________.
   a. ¿Qué va?      b. ¿Cómo es?     c. ¿Quién es?       d. ¿Qué tal?

9. ¡Qué ruido había con los gritos de los niños y el ________ de los perros!
   a. olor          b. sueño         c. hambre          d. ladrar

10. Para saber la hora, don Juan miró el ____________.
a. calendario      b. bolsillo      c. estante       d. despertador

11. Yo, que comprendo poco de mecánica, sé que el auto no puede funcionar sin ___.
   a. permiso      b. comer       c. aceite          d. bocina

12. Nos dijo mamá que era hora de comer y por eso __________.
   a. fuimos a nadar   b. tomamos asiento   c. comenzamos a fumar
   d. nos acostamos pronto

13. ¡Cuidado con ese cuchillo o vas a __________ el dedo!
   a. cortarte       b. torcerte       c. comerte        d. quemarte

14. Tuvo tanto miedo de caerse que se negó a __________ con nosotros.
   a. almorzar       b. charlar       c. cantar       d. patinar

15. Abrió la ventana y miró: en efecto, grandes lenguas de ______ salían llameando de las casas.
   a. zorros       b. serpientes       c. cuero       d. fuego

16. Compró ejemplares de todos los diarios pero en vano. No halló __________.
   a. los diez centavos   b. el periódico perdido   c. la noticia que deseaba
   d. los ejemplos

17. Por varias semanas acudieron colegas del difunto profesor a ______ el dolor de la viuda.
   a. aliviar       b. dulcificar       c. embromar       d. estorbar

18. Sus amigos pudieron haberlo salvado pero lo dejaron __________.
   a. ganar       b. parecer       c. perecer       d. acabar

19. Al salir de la misa me sentía tan caritativo que no pude menos que _______ a un pobre mendigo que había allí sentado.
   a. pegarle       b. darle una limosna       c. echar una mirada       d. maldecir

20. Al lado de la Plaza de Armas había dos limosneros pidiendo ________.
   a. pedazos       b. paz       c. monedas       d. escopetas

21. Siempre maltratado por los niños, el perro no podía acostumbrarse a ______ de sus nuevos amos.
a. las caricias       b. los engaños       c. las locuras       d. los golpes

22. ¿Dónde estará mi cartera? La dejé aquí mismo hace poco y parece que el necio de mi hermano ha vuelto a _______.
   a. dejármela    b. deshacérmela   c. escondérmela   d. acabármela

23. Permaneció un gran rato abstraído, los ojos clavados en el fogón y el pensamiento _______ _______.
   a. en el bolsillo   b. en el fuego   c. lleno de alboroto   d. Dios sabe dónde

24. En vez de dirigir el tráfico estabas charlando, así que tú mismo ___________ del choque.
   a. sabes la gravedad   b. eres testigo   c. tuviste la culpa   d. conociste a las víctimas

25. Posee esta tierra un clima tan propio para la agricultura como para _______.
   a. la construcción de trampas   b. el fomento de motines   c. el costo de vida   d. la cría de reses

26. Aficionado leal de obras teatrales, Juan se entristeció al saber ___________ gran actor.
   a. del fallecimiento   b. del éxito   c. de la buena suerte   d. de la alabanza

27. Se reunieron a menudo para efectuar un tratado pero no pudieron _________.
   a. desavenirse   b. echarlo a un lado   c. rechazarlo   d. llevarlo a cabo

28. Se negaron a embarcarse porque tenían miedo de _________.
   a. los peces   b. los naufragios   c. los faros   d. las playas

29. La mujer no aprobó el cambió de domicilio pues no le gustaba ___________.
   a. el callejeo   b. el puente   c. esa estación   d. aquel barrio

30. Era el único que tenía algo que comer pero se negó a ___________.
   a. hojearlo   b. ponérselo   c. conservarlo   d. repartirlo

**Cloze Test**

In the following text, some of the words have been replaced by blanks numbered 1 through 20. First, read the complete text in order to understand it. Then reread it and choose the correct word to fill each blank from
El sueño de Joan Miró

Hoy se inaugura en Palma de Mallorca la Fundación y Joan Miró, en el mismo lugar en donde el artista vivió sus últimos treinta y cinco años. El sueño de Joan Miró se ha _______ ______ (1). Los fondos donados a la ciudad por el pintor y su esposa en 1981 permitieron que el sueño se _______ ______ (2); más tarde, en 1986, el Ayuntamiento de Palma de Mallorca decidió _______ ______ (3) al arquitecto Rafael Moneo un edificio que _______ ______ (4) a la vez como sede de la entidad y como museo moderno. El proyecto ha tenido que _______ ______ (5) múltiples obstáculos de carácter administrativo. Miró, coincidiendo _______ ______ (6) los deseos de toda su familia, quiso que su obra no quedara expuesta en ampulosos panteones de arte o en _______ ______ (7) de coleccionistas acaudalados; por ello, en 1981, creó la fundación mallorquina. Y cuando estaba _______ ______ (8) punto de morir, donó terrenos y edificios, así como las obras de arte que en ellos _______ ______ (9).

El edificio que ha construido Rafael Moneo se enmarca en _______ ______ (10) se denomina “Territorio Miró”, espacio en el que se han _______ ______ (11) de situar los distintos edificios que constituyen la herencia del pintor.

El acceso a los mismos quedará _______ ______ (12) para evitar el deterioro de las obras. Por otra parte, se _______ ______ (13), en los talleres de grabado y litografía, cursos _______ ______ (14) las distintas técnicas de estampación. Estos talleres también se cederán periódicamente a distintos artistas contemporáneos, _______ ______ (15) se busca que el “Territorio Miró” _______ ______ (16) un centro vivo de creación y difusión del arte a todos los niveles.

La entrada costará 500 pesetas y las previsiones dadas a conocer ayer aspiran _______ ______ (17) que el centro acoja a unos 150.000 visitantes al año. Los responsables esperan que la institución funcione a _______ ______ (18) rendimiento a principios de la _______ ______ (20) semana, si bien el catálogo completo de las obras de la Fundación Pilar y Joan Miró no estará listo hasta dentro de dos años.
Cloze Test Answer Sheet

1. a. cumplido  b. completado  c. terminado
2. a. inició  b. iniciara  c. iniciaba
3. a. encargar  b. pedir  c. mandar
4. a. hubiera servido  b. haya servido  c. sirviera
5. a. superar  b. enfrentarse  c. acabar
6. a. por  b. en  c. con
7. a. voluntad  b. poder  c. favor
8. a. al  b. en  c. a
9. a. habría  b. había  c. hubo
10. a. que  b. el que  c. lo que
11. a. pretendido  b. tratado  c. intentado
12. a. disminuido  b. escaso  c. restringido
13. a. darán  b. enseñarán  c. dirán
14. a. sobre  b. en  c. para
15. a. ya  b. así  c. para
16. a. será  b. sea  c. es
17. a. casos  b. aspectos  c. niveles
18. a. a  b. de  c. para
19. a. total  b. pleno  c. entero
20. a. siguiente  b. próxima  c. pasada

Answer Key: Multiple Choice Test
1. d  
2. a  
3. a  
4. d  
5. c  
6. a  
7. a  
8. d  
9. d  
10. d  
11. c  
12. b  
13. a  
14. d  
15. d  
16. c  
17. a  
18. c  
19. b  
20. c  
21. a  
22. c  
23. d  
24. c  
25. d  
26. a  
27. d  
28. b  
29. d  
30. d

**Answer Key: Cloze Test**

1. a  
2. b  
3. a  
4. c  
5. a  
6. c  
7. b  
8. c  
9. b  
10. c  
11. b  
12. c  
13. b  
14. a  
15. b  
16. b  
17. c  
18. a  
19. b  
20. b  

Total points possible: 50

**Scores:**

- **Advanced** 40 to 50
- **Intermediate** 30 to 39
- **Low** 0 to 29
APPENDIX C: Controlled Vocabulary Levels Productive Abilities Test

Name: ________________________

Instructions: There are 30 questions. Fill in the blank to complete the word. Correct spelling is not required. The example has been done for you.

Example: He was riding a bicycle.

1. I'm glad we had this opp_______ to talk.
2. He has a successful car_______ as a lawyer.
3. Soldiers usually swear an oath_______ of loyalty to their country.
4. Sudden noises at night scare_______ me a lot.
5. There are a dozen_______ eggs in the basket.
6. Ann introduced her boyfriend to her mother.
7. The voter placed the ball_______ in the box.
8. The prisoner was released on parole_______.
9. Every working person must pay income tax_______.
10. To improve the country's economy, the government decided on economic reform_______.
11. They keep their valuables in a vault_______ at the bank.
12. He wasn't very popular_______ when he was a teenager, but he has many friends now.
13. The pirates buried the treasure_______ on a desert island.
14. She wore a beautiful green gown_______ to the ball.
15. A bird perched at the window led_______.
16. The rich man died and left all his wealth_______ to his son.
17. Many gardens are full of flowers.
18. Many people are injured_______ in road accidents every year.
19. The government tried to protect the country's industry by reducing the import_______ of cheap goods.
20. The kitten is playing with a ball of yarn_______.
21. This year long socks_______ are fashionable again.
22. The children's games were amusing at first, but finally got on the parents' nerves_______.
23. The thieves have forced an entry_______ into the building.
24. They had to climb_______ a steep mountain to reach the cabin.
25. He received many com__________ on his dancing skill.
26. After finishing his degree, he entered a new ph__________ in his career.
27. He was on his knees, ple__________ for mercy.
28. The nu__________ was helping the doctor in the operating room.
29. She changes partners often because she cannot have a sta__________ relationship with one person.
30. The prisoner was put in soli__________ confinement.
APPENDIX D: Association Task


cuestion léxica

Instrucciones:
• Escriba todas las palabras que pueda sobre el tema propuesto, de 1 en adelante.
• Es importante no hablar con otros durante la producción de palabras.
• Escriba sin tener en cuenta la ortografía. Aunque tenga dudas, escriba la palabra.
• No se preocupe si la palabra le resulta malsuena o vulgar, escribala igual.
• No escriba nombres propios.
• Tiene solo dos minutos para cada tema.
• Trabaje con un solo tema. No cambie de página.
• Espere la señal de la investigadora antes de pasar al siguiente tema.
• Si necesita más espacio, escriba en la parte en blanco de la hoja.

La ropa

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>21.</td>
</tr>
<tr>
<td>2.</td>
<td>22.</td>
</tr>
<tr>
<td>3.</td>
<td>23.</td>
</tr>
<tr>
<td>4.</td>
<td>24.</td>
</tr>
<tr>
<td>5.</td>
<td>25.</td>
</tr>
<tr>
<td>7.</td>
<td>27.</td>
</tr>
<tr>
<td>8.</td>
<td>28.</td>
</tr>
<tr>
<td>9.</td>
<td>29.</td>
</tr>
<tr>
<td>10.</td>
<td>30.</td>
</tr>
<tr>
<td>11.</td>
<td>31.</td>
</tr>
<tr>
<td>12.</td>
<td>32.</td>
</tr>
<tr>
<td>13.</td>
<td>33.</td>
</tr>
<tr>
<td>14.</td>
<td>34.</td>
</tr>
<tr>
<td>15.</td>
<td>35.</td>
</tr>
<tr>
<td>16.</td>
<td>36.</td>
</tr>
<tr>
<td>17.</td>
<td>37.</td>
</tr>
<tr>
<td>18.</td>
<td>38.</td>
</tr>
<tr>
<td>19.</td>
<td>39.</td>
</tr>
<tr>
<td>20.</td>
<td>40.</td>
</tr>
</tbody>
</table>
VITA
Maryann Parada

EDUCATION

Ph.D.  University of Illinois at Chicago
Hispanic Linguistics, expected: May 2016
Dissertation title: *Lexical availability in diaspora Spanish: An intergenerational analysis of trilingual Chilean Swedes*
Advisor: Dr. Kim Potowski

M.A.  Brigham Young University
Hispanic Linguistics, 2008
Thesis title: *Spanish and Swedish in contact: Transfer in adjective placement*
Thesis director: Dr. Robert Smead

B.A.  Brigham Young University
Double major: Spanish and Linguistics, 2006

Certificates:  Health Care Interpreting (English/Spanish), Waubonsee College, May 2016

Research
Concentrations: sociolinguistics; heritage language studies

Areas of specialization/interest: sociolinguistics of minority Spanish; Spanish heritage learners and pedagogies; heritage language maintenance and development; U.S. Latino/a identity; language maintenance and shift; language contact; language policy and ideology; Latino media; style variation; onomastics; Chilean diaspora

TEACHING

2008 – present  Instructor of Record, Department of Hispanic and Italian Studies, University of Illinois at Chicago

Basic Language Program
1st-4th semester Spanish

Heritage Language Program
Spanish for Heritage Speakers I & II

Advanced Undergraduate courses
Spanish Grammar in Practice
Spanish in the United States

2009 & 2012  Adjunct Instructor, General Education, Kendall College
Professional Spanish I & II
2007  High School Spanish Teacher, **High Bluff Academy**, San Diego, CA

2004-2008  Instructor of Record, Department of Spanish and Portuguese, **Brigham Young University**, 1st-5th semester Spanish

**PUBLICATIONS**

*Published or To Appear*


**In Preparation**

National news media and the standardization of U.S. Spanish  (In preparation for submission to *Multilingua: Journal of Cross-Cultural and Interlanguage Communication*)

Variable adjective placement in trilingual Spanish heritage speakers (In preparation for submission to *Heritage Language Journal*)

Personal names as a stylistic resource: Socio-onomastic perspectives of Spanish receptive bilinguals (In preparation for submission to *Journal of Sociolinguistics*)

Ethnolinguistic emblems in Latino Chicago: Attitudes of the second generation toward names and naming. (In preparation for submission to *Journal of Multilingual and Multicultural Development*)

**PROFESSIONAL PRESENTATIONS**

"How domain-specific is heritage speaker lexicon?: Trilingual Spanish heritage speakers in Sweden."


"Ethnic personal names and Spanish receptive bilinguals: Attitudes and sociolinguistic functions." 24th Conference on Spanish in the U.S. & 9th Conference on Spanish in Contact with Other Languages, McAllen, TX, 2013.


"First-born advantages: Linking birth order, name ethnicity and heritage language proficiency." 23rd Conference on Spanish in the U.S. & 8th Conference on Spanish in Contact with Other Languages, Sacramento, CA, 2011.


"First names and their relationship to the Spanish proficiency and ethnic identity of bilingual university students." In/Between Conference: Thoughts on Literature and Language, University of Illinois at Chicago, 2010.

"Name ethnicity, identity and Spanish proficiency." Fifth International Workshop on Spanish Sociolinguistics, North Carolina State University, 2010.

"An online placement exam for Spanish heritage speakers and L2 students." First International Conference on Heritage Languages, UCLA (With Kim Potowski), 2010.

"Spanish and Swedish in contact: Transfer in adjective placement." 22nd Conference on Spanish in the U.S. & 7th Conference on Spanish in Contact with Other Languages, Miami, 2009.

EDITORIAL EXPERIENCE
2013-present  Editorial Assistant, *Spanish in Context* (John Benjamins)


**OTHER RESEARCH / PROFESSIONAL EXPERIENCE**

2013-2015  Member, University of Illinois-Chicago Language in Context Research Group

2009-2010  Project Team Member (with Professors Kim Potowski and Kara Morgan–Short)  
Project: Developing a universal Spanish placement exam for heritage and L2 learners  
Department of Hispanic and Italian Studies, UIC

2005-2006  Research Assistant to Dr. Jerry Larson, Professor of Spanish  
Project: Developing an online context-based diagnostic Spanish exam  
Humanities Technology and Research Support Center, BYU

**AWARDS & HONORS**

2016  Emerging Scholar Award, American Name Society

2016  Chancellor’s Graduate Research Fellowship, University of Illinois-Chicago

2013  President’s Research in Diversity Award, University of Illinois

2012  Provost’s Award, University of Illinois-Chicago

2011  Grantee, National Heritage Language Resource Center

2008  Outstanding Graduate Student Instructor of the Year  
BYU, Department of Spanish and Portuguese

**UNIVERSITY AND PROFESSIONAL SERVICE**

Webmaster for Talks in Linguistics (UIC), 2014-2015  
Graduate Student Council departmental representative (UIC), 2014-2015  
Graduate assistant for New Ways of Analyzing Variation (N WAV) conference, Chicago, 2014  
Spanish placement exam consultant (UIC), 2010-2012  
Committee member for annual Bilingualism Forum (UIC), 2010-2012  
Community translation service, Chicago Campaign for Better Health Care, 2009  
Vice President of Graduate Association of Spanish and Portuguese (BYU), 2006-2007  
Secretary of Sigma Delta Pi, BYU Chapter, 2007-2008  
Student host for visiting scholars (Humberto López Morales, José Antonio Samper Padilla & Clara Hernández) (BYU), 2008  
Volunteer at annual “Españolandia” high school student language fair (BYU), 2005-2008
**GRADUATE COURSEWORK**

**Language Pedagogy**
- Directed Teaching of Spanish (BYU)
- Testing Language Skills (BYU)
- Heritage Learner Writing (UIC)

**Advanced Grammar**
- Problems in Spanish Grammar (BYU)
- Spanish Morphosyntax (BYU)
- Phonetics and Phonology (BYU)
- Advanced Syntax (UIC)

**Bilingualism / Language Contact**
- Bilingualism (BYU)
- Border Spanish (BYU)
- History of the Spanish Language (BYU)
- Second Language Acquisition and Bilingualism (UIC)

**Sociolinguistics / Dialectology**
- Spanish Sociolinguistics (BYU)
- Hispanic Dialectology (BYU)
- Language and Gender (UIC)
- Violence and Language (UIC)
- Language Policy and Cultural Identity (UIC)
- History of European Standard Languages (UIC)

**Literature**
- Hispanic American Short Story (BYU)
- Golden Age Literature (BYU)
- Spanish Theatre Production (BYU)

**Summer Workshops**

Heritage Language Research Institute, *National Heritage Language Resource Center*  

**PROFESSIONAL MEMBERSHIPS**

American Association of Teachers of Spanish and Portuguese (AATSP)  
American Name Society (ANS)  
Sigma Delta Pi

**LANGUAGES**

<table>
<thead>
<tr>
<th>Language</th>
<th>Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>Near Native</td>
</tr>
<tr>
<td>French</td>
<td>Reading ability</td>
</tr>
<tr>
<td>Swedish</td>
<td>Intermediate</td>
</tr>
<tr>
<td>Arabic</td>
<td>Beginner</td>
</tr>
<tr>
<td>Portuguese</td>
<td>Reading ability</td>
</tr>
</tbody>
</table>
REFERENCES

Kim Potowski
Associate Professor
Department of Hispanic and Italian Studies,
University of Illinois at Chicago
kimpotow@uic.edu
(312) 996-8524

Kara Morgan-Short
Associate Professor
Department of Hispanic and Italian Studies,
University of Illinois at Chicago
karams@uic.edu
(312) 996-5743

Richard Cameron
Associate Professor
Department of Linguistics
University of Illinois at Chicago
rcameron@uic.edu
(312) 996-3241