

THE EFFECT OF EXPOSURE TO PHONOLOGICAL VARIATION ON PERCEPTUAL CATEGORIZATION
AND LEXICAL ACCESS IN SECOND LANGUAGE SPANISH: THE CASE OF /S/-ASPIRATION IN
WESTERN ANDALUSIAN SPANISH

Robert William Bedinghaus

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Spanish and Portuguese,
Indiana University
June 2015

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Doctoral Committee

Kimberly Geeslin, Ph.D.
Chairperson

Isabelle Darcy, Ph.D.

Erik Willis, Ph.D.

Manuel Díaz-Campos, Ph.D.

May 4, 2015

To my wife, Kate Bedinghaus

with immense thankfulness for her unwavering support, help, encouragement, and for making our home a wonderful place to return to after the many long days of work that it took to complete this degree, and in particular this dissertation.

Acknowledgements

Before proceeding any further I want to thank my Lord and Savior Jesus Christ, who, while I was yet a sinner, made me new through His own sacrifice (Romans 5:8) and gives me strength each day for the work He has given me to do in this world. I also want to thank the pastors, elders, and friends at my local church, Clearnote Church Bloomington, for their faithful work in proclaiming the truth of the gospel of Jesus Christ that has strengthened me and changed me for the better in many ways during my time in Bloomington.

Second, I would like to thank my beautiful and excellent wife, Kate. Through the ups and downs of graduate school she has remained a solid rock for me, supporting my work, encouraging me, and helping me to keep things in perspective. Furthermore, she is a very hard worker in our home, caring for our boys and making our home a wonderful place to be. It is such a sweet balm for me to get home at the end of the day and have my wife and sons smiling at me and welcoming me cheerfully. The difficulties of the work day fade away and I am reminded that my work is much more than the research and teaching I have been doing for the past six years. I also want to thank my sons, Liam (4 1/2) and Judah (20 months) for the constant laughter, energy, fun, and love that they bring to our family. They have kept graduate school from taking over my life!

Next I would like to thank my dissertation director Dr. Kimberly Geeslin, and my committee members Dr. Isabelle Darcy, Dr. Erik Willis, and Dr. Manuel Diaz-Campos. All of them are excellent at what they do and have brought different and very helpful perspectives on my work. I am deeply grateful for the time and effort that they have invested in me. Dr. Geeslin always amazes me with the amount of work she can fit into a day, and the speed and quality of her feedback as I turned in chapter after chapter to her made this dissertation much better. The feedback of the other committee members along the way has also benefitted me and my work greatly.

I also want to thank others who have made this dissertation possible. I want to thank the directors of the two study abroad programs in Seville, Leslie Routman, Jaime Ramírez, and Antonio Rodríguez, for letting me attend classes to recruit their students and collect data in their computer labs. Thank you, also, to the faculty and staff of both programs for accommodating my research, participating, and providing connections to recruit native Spanish speakers to participate in my study. Thank you to my good friends in Seville (Juan Manuel, Nubia, Enrique) who housed me, often fed me, spent time with me, and helped me recruit native speakers for my study. I also thank the instructors at IU who let me recruit students from their Spanish classes, and a huge thanks to the students and native speakers both in Seville and at IU who took time to participate in my study. I could not have done it without you! Thank you to Bret Linford for helping me collect data at IU while I was in Spain. I also wish to thank Wes Beaulieu, Kate, and Bonnie at the IU Statistical Consulting Center for their invaluable help in analyzing my data. Thank you to Natasha Branch in the recording studio in Ballantine Hall for helping me record all of my stimuli for my speech perception tasks, and for the three native speakers who put up with reading a lot of words (and fake words) into a microphone. I would also like to thank my colleagues, who have always encouraged me and are very professional, hard-working individuals who will make great professors.

And finally, I want to thank my family members who have supported me throughout this process. Thank you, especially, to my parents Bob and Lori Bedinghaus, for always being there for me in innumerable ways and supporting us through this graduate school journey. Thank you for raising me well and for always encouraging me. Thank you to my sisters, Lauren and Kayla, for their love and support. Thank you to my in-laws, Royce and Marci Yoder, who have loved and accepted me into their family and allowed me to marry their wonderful daughter. I'm blessed to have such wonderful in-laws!

Robert William Bedinghaus

THE EFFECT OF EXPOSURE TO PHONOLOGICAL VARIATION ON PERCEPTUAL CATEGORIZATION
AND LEXICAL ACCESS IN SECOND LANGUAGE SPANISH: THE CASE OF /S/-ASPIRATION IN
WESTERN ANDALUSIAN SPANISH

This dissertation investigates the acquisition of a dialect-specific phonological variant by American English-speaking L2 learners of Spanish who study abroad for one semester in the target dialect region compared to L2 learners who remain at their home university in the traditional language classroom context ('At-home' learners) without exposure to the target variant. The perception of aspirated-/s/ in Western Andalusian Spanish, which is unique in its phonetic articulation and is subject to social and stylistic variation, is examined. In this investigation, 33 native speakers of Spanish from Seville, Spain, 10 native speakers of Spanish from regions in which /s/ is not aspirated, 48 L2 learners of Spanish in the study abroad context, and 25 learners of Spanish in the traditional classroom context completed a forced-choice identification task, a lexical decision task, background questionnaires, a grammar test, and a word familiarity questionnaire. Analysis of the patterns of identification and lexical decision related to /s/-aspiration revealed: (1) effects of *condition* (i.e., the aspirated condition was significantly more difficult for L2 learners than other control conditions), (2) effects of *exposure* to /s/-aspiration for the L2 groups on the accuracy of identification and lexical decision (i.e., study abroad groups made gains while the at-home group did not), (3) effects of *exposure* to /s/-aspiration on lexical decision response time (i.e., study abroad learners responded faster to /s/-aspiration over time while at-home learners did not), (4) and some effects of the amount and types of *target language use* by the L2 learners abroad for the identification task. However, extralinguistic factors such as contact with native speakers and target language use during the semester abroad did not show consistent effects across tasks, which is argued to be a methodological issue. Implications for the study of the effects of

extralinguistic factors on linguistic gains during study abroad are discussed. Differences were also found between the two native speaker groups, primarily for the /s/-aspiration condition, indicating dialectal differences in L1 speech perception. This dissertation shows that immersion during study abroad leads to perceptual gains related to a dialect-specific phonological variant, but that traditional research methods for investigating individual differences among study abroad learners must be revised.

Table of Contents

Chapter 1: Introduction	1
Context of Learning in Second Language Acquisition	2
Why Context of Learning Matters	2
Extralinguistic Factors that Can Affect Outcomes during Study Abroad	5
Study Abroad Outcomes	15
Chapter 2: Review of the Literature	28
The Acquisition of Variable Linguistic Forms in a Second Language.....	28
Variation in Linguistic Systems.....	28
The Acquisition of Variable Structures in Second Language Spanish	33
The Acquisition of Sociolinguistic Variants in Immersion Contexts.....	37
The Perception of L2 Dialect-Specific Variants	48
Spanish /s/	56
A Phonetic Characterization of Andalusian /s/-aspiration	65
Variation of /s/ in Andalusian Spanish According to Social and Contextual Factors.....	76
Second Language Phonology.....	78
Models of L2 Speech Perception	80
Application of L2 Perception Models and Theories to the Current Study.....	84
Models of Spoken Word Recognition	89
Two Task Paradigms for Measuring the Perception of L2 Phones	92
Forced-Choice Identification	92
Lexical Decision	96
Summary of the Motivations of the Current Study.....	99
Research Questions.....	103
Chapter 3: Method	104
Participants.....	104
Native Speaker Groups	105
L2 Learner Groups.....	109
Instruments	125
General Procedure	125
Forced-Choice Identification Task.....	127
Procedure.....	133
Lexical Decision Task.....	134
Procedure.....	141
Preparation and Selection of Experimental Stimuli	141
Word Familiarity Questionnaire	149
Grammar Test	150
Language Background Questionnaire (Time 1).....	150
Native Speaker Contact and Target Language Use Questionnaire (Time 2).....	151
Native Speaker Background Questionnaire	155
Data Analysis	155
Data Coding.....	155
Preparation of the Data for Analysis	157
Analysis of Accuracy According to Group and Condition (ID)	159
Analysis of Response Type (ID) according to Group and Condition	160
Analysis of the Accuracy according to Group and Condition (LD)	160

Analysis of Lexical Decision Response Time	161
Analyses of the Effect of Time on L2 Learners' Performance (ID/LD)	161
Analyses of Extralinguistic Factors for the Study Abroad Groups (ID/LD)	161
Summary of Chapter 3	162
Chapter 4: Results	163
Results of the Forced-Choice Identification Task	165
ID Accuracy for each Coda Condition according to Group (Time 1)	165
ID Accuracy for each Listener Group according to Coda Condition (Time 1)	168
ID Accuracy for each Coda Condition according to Group (Time 2)	171
ID Accuracy for each Listener Group according to Coda Condition (Time 2)	173
The Effect of Time (i.e., Exposure) on L2 Learners' ID Accuracy	176
Individual Patterns of Study Abroad Participants in the Aspirated Condition	180
Response Types by Coda Condition for Each Group and Time	181
The Effects of Extralinguistic Factors on Study Abroad Learners' ID Accuracy	190
Results of the Lexical Decision Task	199
Lexical Decision Accuracy for each Condition according to Group (Time 1)	200
Lexical Decision Accuracy for each Group according to Condition (Time 1)	203
Lexical Decision Accuracy for each Condition according to Group (Time 2)	206
Lexical Decision Accuracy for each Group according to Condition (Time 2)	209
The Effect of Time (i.e., Exposure) on L2 Learners' LD Patterns.....	212
Individual Patterns of Study Abroad Participants in the Asp Real Word Condition	217
The Effects of Extralinguistic Factors on Study Abroad Learners' LD Patterns.....	218
Analysis of Lexical Decision Response Time (RT).....	224
Results for each Condition according to Group (Time 1).....	226
Results for each Group according to Condition (Time 1).....	230
Results for each Condition according to Group (Time 2).....	233
Results for each Group according to Condition (Time 2).....	238
The Effect of Time (i.e., Exposure) on L2 Learners' RTs.....	241
Summary of Chapter 4.....	248
Chapter 5: Discussion and Conclusions	252
The Perception of Word-Internal /s/-Aspiration in Western Andalusian Spanish by L2 Learners – Research Question #1.....	254
Initial Identification Patterns (Time 1).....	254
Native Speakers	273
L2 Learners	256
Response Types	256
L2 Group Differences.....	259
Summary	263
Initial Lexical Access Patterns (Time 1)	263
Native Speakers	264
L2 Learners	265
L2 Group Differences.....	266
Summary	270
The Effect of Longitudinal Exposure to Western Andalusian Spanish on Orthographic Association and Lexical Access Patterns – Research Question #2	271
Orthographic Association over Time	271
Lexical Access over Time.....	274

Summary	277
The Effect of Extralinguistic Factors on Orthographic Association and Lexical Access Patterns – Research Question #3.....	278
Study Abroad Program.....	278
Grammar Test Score	280
Contact with Native Speakers.....	283
Target Language Use.....	285
Target Language Use and Orthographic Association	285
Target Language Use and Lexical Access.....	288
Summary of Target Language Use Factors for Both Tasks	289
Summary of the Effects of Extralinguistic Factors on Study Abroad Learners’ Perception of /s/- aspiration in Andalusian Spanish	296
Implications for Our Understanding of Study Abroad and Second Language Acquisition.....	299
Implications for Study Abroad Program Design	302
Limitations and Future Directions	304
Conclusions.....	312
REFERENCES.....	319
APPENDIX A: Grammar Test.....	358
APPENDIX B: Example of Identification Task Screenshot	359
APPENDIX C: Acoustic Measurements of the Identification Task Target Stimuli	360
APPENDIX D: Waveforms and Spectrograms of Stimuli Representing the Coda Conditions from the Identification Task	363
APPENDIX E: Acoustic Measurements of the Lexical Decision Task Target Stimuli	367
APPENDIX F: Waveforms and Spectrograms Comparing the Target Lexical Decision Coda Conditions .	372
APPENDIX G: Language Background Questionnaire (Time 1 study abroad questionnaire)	373
APPENDIX H: Language Background Questionnaire (Time 1 At-home questionnaire)	382
APPENDIX I: Native Speaker Contact and Target Language Use Questionnaire (Time 2 study abroad questionnaire)	392
APPENDIX J: Native Speaker Contact and Target Language Use Questionnaire (Time 2 At-home questionnaire)	404
APPENDIX K: Native Speaker Questionnaire	414
APPENDIX L: Word Familiarity Questionnaire	420
APPENDIX M: Scatterplots of Correlations between Language Use Factors and Accuracy in the Aspirated Condition for Both Tasks	427
APPENDIX N: Target language use and Contact with NS descriptives by group (Study Abroad Groups)	431
CURRICULUM VITAE	

List of Tables and Figures

Tables

Table 1: Factors affecting outcomes during study abroad	7
Table 2: Mean durations of VOT and stop closures in Seville and Cádiz for /VS.CV/ and /CV.CV/ sequences (O'Neill, 2010 and Torreira, 2007)	72
Table 3: Statistics for the duration of coda glottal activity (i.e. GLOT) in both cities of Western Andalusia from O'Neill (2010).....	72
Table 4: Statistics for mean VOT in /VS.tv/ contexts compared to /V.tv/ contexts, comparing older to younger speakers in East and West Andalusia (Ruch & Harrington, 2014)	72
Table 5: Statistics for mean stop closure duration in /VS.tv/ contexts compared to /V.tv/ contexts, comparing older to younger speakers in East and West Andalusia (Ruch & Harrington, 2014) 72	
Table 6: Postulates (P) and Hypotheses (H) of Flege's (1995) SLM (p. 239).....	82
Table 7: PAM Categories of Assimilation (Best, 1994, 1995).....	83
Table 8: Summary of PAM and SLM predictions for the perception of /s/-aspiration.....	88
Table 9: Duration of stays outside of Western Andalusia (SEVILLE group)	106
Table 10: Self-reported knowledge of languages other than Spanish by the SEVILLE participants	107
Table 11: Demographic information for Non-Aspirating group.....	108
Table 12: Self-reported knowledge of languages other than Spanish by the NON-ASP participants	109
Table 13: Number of participants with each level of Spanish education prior to data collection (Study Abroad 1 group)	113
Table 14: Number of participants reporting Spanish use in each category prior to study abroad (Study Abroad 1 group)	114
Table 15: Number of participants with each level of Spanish education prior to data collection (Study Abroad 2 group)	118
Table 16: Number of participants reporting Spanish use in each category prior to study abroad (Study Abroad 2 group)	119
Table 17: Number of participants with each level of Spanish education prior to data collection (At-home group).....	121

Table 18: Number of participants with each level of Spanish education prior to data collection (At-home group).....	123
Table 19: Summary of the participant groups	124
Table 20: Examples of each Identification task coda condition	128
Table 21: Balance of preceding vowels and following consonants across versions of the Identification task	129
Table 22: Target stimuli for version one of the Identification task.....	129
Table 23: Target stimuli for version two of the Identification task	130
Table 24: The 50 control stimuli for the Identification task	132
Table 25: The 88 distracter stimuli (with affricate/fricative stimuli by version) for the Identification task	132
Table 26 Example of the experimental conditions of the lexical decision task	135
Table 27: Distribution of vowels preceding /s/, /ptk/ following /s/, and stress of the syllable containing coda /s/ for the real words in both Lexical Decision task versions.....	138
Table 28: CodaS and Asp experimental stimuli for the lexical decision task	139
Table 29: The 108 Distracter stimuli in the Lexical Decision Task	140
Table 30: Descriptive statistics of the stimuli selected for the forced-choice identification task (in milliseconds).....	145
Table 31: VOT statistics according to following consonant and coda condition for the stimuli of the forced-choice identification task (in milliseconds).....	145
Table 32 VOT statistics according to condition for each talker (forced-choice identification task)	146
Table 33: Descriptive statistics for CodaS and Asp target stimuli of the lexical decision task (in milliseconds).....	147
Table 34: Descriptive statistics for VOT according to each voiceless stop and each condition for the lexical decision task (in milliseconds).....	148
Table 35: VOT statistics according to condition for each talker (lexical decision task)	148
Table 36: Background information elicited by the Time 1 background questionnaire	152

Table 37: NS contact and language use information elicited by the Time 2 questionnaire	153
Table 38: Items eliminated from the Lexical Decision task analysis	158
Table 39: Items eliminated from the Identification task analysis	159
Table 40: Overall Identification accuracy (%) by Group and Coda Condition (Time 1).....	166
Table 41: Overall Identification accuracy (%) by Group and Coda Condition (Time 2).....	172
Table 42: Identification accuracy for each L2 group by coda condition and time (T1/T2)	178
Table 43: Individual patterns of according to study abroad group for the Asp condition over time.....	181
Table 44: Distribution of Identification responses for the SEVILLE group according to coda condition	182
Table 45: Distribution of Identification responses for the NON-ASP group according to coda condition	183
Table 46: Distribution of Identification responses for the At-home group according to coda condition (Time 1)	184
Table 47: Distribution of Identification responses for the At-home group according to coda condition (Time 2)	185
Table 48: Distribution of Identification responses for the Study Abroad 1 group according to coda condition (Time 1).....	186
Table 49: Distribution of Identification responses for the Study Abroad 1 group according to coda condition (Time 2).....	187
Table 50: Distribution of Identification responses for the Study Abroad 2 group according to coda condition (Time 1).....	188
Table 51: Distribution of Identification responses for the Study Abroad 2 group according to coda condition (Time 2).....	189
Table 52: Distribution of Identification responses to Asp stimuli for all groups, including Time 1 and Time 2 responses for the L2 groups.....	190
Table 53: Descriptive statistics for the language use factors as reported by the study abroad learners on the questionnaire at Time 2 (Identification task)	193

Table 54: Spearman’s rank correlations for accuracy in the aspirated condition at Time 2 and language use factors (Identification task).....	194
Table 55: Spearman’s rank correlations for Time1 to Time 2 accuracy change over time in the aspirated condition and language use factors (Identification task)	196
Table 56: Mean Lexical Decision accuracy (%) for each Condition by Group for NSs and L2 learners at Time 1.....	201
Table 57: Lexical Decision accuracy (%) for each Condition by Group for NSs and L2 learners at Time 2	207
Table 58: Mean accuracy (%) and standard deviations for each L2 group in each condition at Time 1 and Time 2 (Lexical Decision task)	213
Table 59: Number of participants in each study abroad group that increased, decreased and stayed the same in the Asp real word condition from Time 1 to Time 2	218
Table 60: Descriptive statistics for the language use factors as reported by the study abroad learners on the questionnaire at Time 2	220
Table 61: Spearman’s rank correlations for lexical decision accuracy in the Asp real word condition at Time 2 and language use factors	220
Table 62: Mean RTs (ms) and standard deviations for each Condition according to Group in milliseconds (Time 1)	227
Table 63: Mean log(RT) and standard deviations for each Condition according to Group (Time 1)	228
Table 64: Mean RTs (ms) and standard deviations for each Condition according to Group in milliseconds (Time 2)	234
Table 65: Mean log(RT) and standard deviations for each Condition according to Group (Time 2)	235
Table 66: Comparison of the L2 groups’ RTs (ms) in each Condition according to Time	243
Table 67: Comparison of the L2 groups’ log(RT) in each Condition according to Time.....	244
Table 68: Summary of the results of the Identification task	250
Table 69: Summary of the results of the Lexical Decision task	251
Table 70. Descriptive statistics for the language use factors as reported by the study abroad learners on the questionnaire at Time 2.....	431

Table 71. Descriptive statistics for the Contact with Native Speakers factors as reported by the study abroad learners on the questionnaire at Time 2 432

Figures

Figure 1: Example of a sibilant, *susto* [susto]..... 65

Figure 2: Example of aspiration from Schmidt (2011, p. 11) - *chiste* [tʃihte]..... 66

Figure 3: Example of NOGLOT with a post-aspirated /t/, *puesto* [puet^ho]..... 67

Figure 4: Example of breathy voicing with a post-aspirated /t/, *susto* [sut^ho]..... 68

Figure 5: Example of a regular /t/ (not post-aspirated), *susto* [susto] (14.8 ms VOT) 69

Figure 6: Example of a post-aspirated /t/, *susto* [sutho] (51.8 ms VOT) 69

Figure 7: Summary of mean Identification accuracy by Group Coda Condition at Time 1 (N=116)..... 170

Figure 8: Summary of mean Identification accuracy by Coda Condition – All groups at Time 2 (N=116) 176

Figure 9: Accuracy in the Asp condition of the identification task over time (L2 groups)..... 180

Figure 10: Summary of Mean Lexical Decision accuracy (%) for each Condition according to Group and Lexical Status (Time 1)..... 206

Figure 11: Summary of Mean Lexical Decision Accuracy for each Group according to Condition and Lexical Status (Time 2) 211

Figure 12: Mean Lexical Decision accuracy by Condition and Time for the Study Abroad 1 group (N=23)..... 214

Figure 13: Mean Lexical Decision accuracy by Condition and Time for the Study Abroad 2 group (N=25)..... 215

Figure 14: Mean Lexical Decision accuracy by Condition and Time for the At-home group (N=25)..... 216

Figure 15: Summary of the Lexical Decision accuracy results for the L2 groups according to Time and Condition (N=73) 217

Figure 16: Summary of Response Time for Condition according to Group (Time 1)..... 230

Figure 17: Summary of Response Time for Condition according to Group (Time 2)..... 238

Figure 18: Response Time for each Group according to Condition at Time 2 (N=116) 240

Figure 19: Comparison of the Study Abroad 1 group’s mean Response Time in each Condition according to Time	245
Figure 20: Comparison of the Study Abroad 2 group’s mean Response Time in each Condition according to Time	246
Figure 21: Comparison of the At-home group’s mean Response Time in each Condition according to Time	247
Figure 22: Summary of the Lexical Decision accuracy results for the L2 groups according to Time and Condition (N=73).....	276
Figure 23. Praat screenshot of the identification response options for the stimulus <gasco> ([gak ^h o]).....	359
Figure 24: Post-aspiration, <i>jasta</i> [jat ^h a]	363
Figure 25: Sibilant [s], <i>fesco</i> [fesko]	363
Figure 26: No coda, <i>jata</i> [xata]	364
Figure 27: [f], <i>tifpe</i> [tifpe]	364
Figure 28: [l], <i>galca</i> [galka]	365
Figure 29: [n], <i>binco</i> [biŋko].....	365
Figure 30: [r], <i>narto</i> [narto].....	366
Figure 31: Example of a CodaS condition, <i>sosca</i> [soska]	372
Figure 32: Example of the Asp condition, <i>sosca</i> [sok ^h a]	372
Figure 33. Correlation between mean ID accuracy at Time 2 in the aspirated condition and the reported percentage of the average day spent speaking in Spanish with NS of Spanish	427
Figure 34. Correlation between mean ID accuracy at Time 2 in the aspirated condition and the reported percent use of Spanish overall.....	427
Figure 35. Correlation between mean ID accuracy at Time 2 in the aspirated condition and the reported percent use of Spanish while writing	428
Figure 36. Correlation between T1/T2 ID accuracy differential (aspirated) and the reported percentage of the average day spent speaking Spanish with NSs of Spanish	428
Figure 37. Correlation between T1/T2 ID accuracy differential (aspirated) and the reported percent use of Spanish while writing.....	429

Figure 38. Correlation between mean lexical decision accuracy and the percent of an average day spent watching media in English..... 429

Figure 39. Correlation between mean lexical decision accuracy and the percent of an average day spent speaking in English with NSs of Spanish..... 430

Figure 40. Correlation between mean lexical decision accuracy and the percent of an average day spent reading in Spanish 430

Conventions

L1 – First language

L2 – Second language

SA – Study Abroad

AH – At-home group (i.e., L2 learners in the traditional university classroom learning context)

NON-ASP – Non-Aspirating group (i.e., native speakers from regions of the Spanish speaking world where /s/ is not aspirated)

NS – Native speaker

NNS – Non-native speaker

WAS – Western Andalusian Spanish

EAS – Eastern Andalusian Spanish

ID - Forced-Choice Identification Task

LD – Lexical Decision Task

Chapter 1 - Introduction

This dissertation is grounded in theories and methods of second language acquisition and second language phonology research in order to investigate the acquisition of the perception of a dialect-specific phonological variant that is both geographically and socially variable: the aspiration of /s/ in Western Andalusian Spanish. The goal of this study is to contribute to our understanding of the acquisition of variable linguistic forms in a second language, particularly the acquisition of phonological variation under conditions of exposure in a study abroad context. Such studies are crucial to SLA theory because, as Tarone (2007) argued, the only real way to determine whether social context plays a role in second language acquisition is to conduct research that follows L2 learners' development of specific L2 forms over time in view of their interactions with others in specific social contexts, and that conducting research in the study abroad context is a promising way of doing so (p. 845). Therefore, the current study employs a longitudinal methodology that investigates two groups of L2 learners who studied abroad in Seville, Spain for one semester at one of two different study abroad programs in comparison to a group of L2 learners that did not study abroad during the same 10- to 11-week time period. The L2 learners were tested on their perception of /s/-aspiration in Andalusian Spanish by means of two speech perception tasks: forced-choice identification, which tests how the learners categorize /s/-aspiration in terms of orthographic representation, and lexical decision, which measures the speed and accuracy with which the learners are able to access words from the mental lexicon that contain /s/-aspiration compared to words that do not contain this phonological variant. The study design also includes an analysis of learners in the study abroad context to determine whether the extralinguistic factors of contact with native speakers, target language use, grammatical proficiency and study abroad program predict gains in performance on identification and lexical access measures over time. Additionally, two groups of native speakers serve as a baseline for comparison to the L2 learners: a group of native

speakers from Seville, Spain, and a group of native speakers from other parts of the Spanish-speaking world where /s/ is not aspirated.

This chapter begins with a description of the importance to the field of second language acquisition (SLA) of research that focuses on the context in which L2 learners acquire a second language. This will be followed by a review of general findings related to the acquisition of a second language in different contexts of learning, including extralinguistic factors have been shown to play a role in language acquisition during study abroad, as well as results of studies on the acquisition of different types of linguistic structures in different contexts of learning. The result of this review will be a clear gap in the literature on context of learning and how the current study begins to fill that gap.

Context of Learning in Second Language Acquisition

Why Context of Learning Matters in SLA

Before reviewing research that has focused on context of learning and how it is related to second language acquisition, it is necessary to understand why we should consider the context in which someone acquires the second language in the first place. The context in which a learner acquires a second language has become an important focus of research on second language acquisition for two primary reasons. The first reason is that the quality and quantity of input are crucial to the second language acquisition process (Krashen, 1982; Long, 1981; VanPatten, 2004). SLA theories have focused strongly on the importance of input and learners' interactions with people. For example, Long's Interaction Hypothesis (1996) emphasized the negotiation of information breakdowns through interaction in order to foment acquisition. Different learning contexts, such as a traditional classroom at a home university, an intensive immersion program either in the home country or abroad, or a study abroad program, can provide to the L2 learner different types and amounts of input, as well as different opportunities for the negotiation of linguistic meaning (Howard, 2011; Tarone, 2000; Tarone & Swain,

1995). For example, in the traditional classroom (henceforth At-home, or “AH” context) the linguistic input to which a learner is exposed is typically limited to learner-teacher or learner-learner interactions. As Lafford (2006) notes, apart from rare opportunities for access to authentic input through materials such as online videos, there are relatively few opportunities for learners in a classroom environment to experience grammar and vocabulary in a meaningful way. Furthermore, she notes that the processing of input in the classroom setting is often limited to the word and sentence levels and rarely reaches more complex discourse level processing. These are important input issues that can differ from the study abroad (henceforth SA) context, where input has more potential to be complex and meaningful. In the study abroad context, a language learner has the opportunity to interact with native speaker (NS) instructors, host families, friends, and other members of the target language community such as employees of establishments that they visit (Lafford, 2006). Nevertheless, the mere availability of more opportunities for meaningful input does not necessarily lead to use of those opportunities and acquisition, since the study abroad context is not always as input-rich as is often believed (Diao, Freed, & Smith, 2011).

This leads to the second reason that context of learning is important to SLA, which is that the research methods of the field as a whole have been increasingly expanding to account for individual, sociocultural and contextual factors that might affect access to input and the processing of input, and consequently second language acquisition (Firth & Wagner, 2007; Tarone, 2000). Some SLA theorists have argued that social context has no bearing on SLA and that SLA is a strictly cognitive phenomenon (e.g., Gregg, 1990; Long, 1997, 1998). However, recent research has shown that it is necessary to account for factors that might affect SLA such as the amount and types of contact a learner has with native speakers, the social characteristics of learners and interlocutors, and learner attitudes toward the language or a specific dialect or dialectal feature. For instance, some research has shown that social and individual factors, such as the sex of the language learner, ethnicity, and attitudinal or motivational

factors, can potentially limit access to input (Brecht, Davidson, & Ginsburg, 1995; Howard, 2011; Gatbonton & Trofimovich, 2008). As has been seen in recent research on the acquisition of sociolinguistic variation (see Geeslin, 2011b for a review), exposure to variable linguistic forms in the input appears to be necessary, though not sufficient in and of itself, for the acquisition of these forms (Regan, Howard, & Lemée, 2009; Thomas, 2004). In other words, since learners have different opportunities for input and make different use of opportunities that they have for input, which results in different outcomes, studying which individual factors lead to more or less opportunities for input and use of that input by the learner is crucial to understanding the acquisition of variable structures, particularly in the study abroad context.

An example can be seen in Segalowitz, Freed, Collentine, Lafford, Lazar and Díaz-Campos (2004), which reviews multiple studies from a special issue in *Studies in Second Language Acquisition* in which an at-home group in the United States and a study abroad group in Spain were tested on multiple aspects of L2 linguistic competence, including pronunciation, grammatical and lexical competencies, communication strategies, and fluency. In conjunction with their experiments all participants completed a detailed Language Contact Profile (LCP: Freed, Dewey, Segalowitz, & Halter, 2004), a questionnaire that elicited information regarding the learners' L2 background prior to study abroad and also their contact with the target language and their own use of the target language during their time abroad. Through the LCP they were able to determine the effects of factors such as contact time with native speakers, passive (e.g., watching television in the target language) and active (e.g., interaction with NSs) target language use, and language learning experience prior to study abroad on the learners' development, giving a more holistic view of factors affecting development in different learning contexts. They were able to show that SLA in the study abroad context does not depend solely on whether a learner studied abroad or not, but that the learners' contact with native speakers and their language use during the semester abroad played a role to differing degrees depending on which aspect of the L2

linguistic system was being studied, since the various linguistic competencies (e.g., grammar, pronunciation, vocabulary) tested were affected differently by the individual factors included in the analysis. Findings such as these point to the necessity of investigating context of learning using a multi-faceted approach that takes into account factors that differentiate individual learners during study abroad rather than a simple study abroad versus at home comparison based on the inaccurate assumption that simply being in the target language culture automatically leads to acquisition. Recent research has been able to shed light on important differences between learning contexts and factors that can help or hinder learners in the acquisition process.

Extralinguistic Factors that Can Affect Outcomes during Study Abroad

Since input is essential to second language acquisition and different contexts of learning can provide different opportunities for input to L2 learners, research on context of learning has investigated how the outcomes of a study abroad experience can vary according to different individual, sociolinguistic, and contextual factors that can affect learners' exposure to and processing of input. A summary of such factors is presented in Table 1. For the current study, the four factors that are of most interest are contact with native speakers and target language use during SA, whether different study abroad programs contribute to different outcomes, and language proficiency level (i.e., grammatical proficiency). Of the factors that have been studied to date, contact with native speakers and target language use by the L2 learners have been the two factors most often incorporated into study designs through the use of questionnaires. Grammatical proficiency is most often measured by means of some form of grammar test.

Grammatical proficiency. One extralinguistic factor of interest to the current study is how grammatical proficiency does or does not predict gains over time in the study abroad context. This is important because the study abroad context presents a lot of complexity to the language learner and

their preparation before entering the study abroad context is considered to be crucial to determining how L2 learners make use of opportunities for input during study abroad and predicting linguistic outcomes (Dekeyser, 2010; Dewey, Bown, Baker, Martinsen, Gold & Eggett, 2014; Magnan and Back, 2007; Segalowitz, Freed, Collentine, Lafford, Lazar & Díaz-Campos, 2004). Nevertheless, the results of studies to date on the connection between grammatical proficiency and gains made during study abroad have been conflicting by showing that the connection between the two likely depends on what aspect of linguistic competence is being observed by researchers, how it is measured, and how grammatical proficiency is measured. For example, Brecht, Davidson and Ginsberg (1995) found that L2 learners with higher grammatical and reading proficiency at the beginning of study abroad, as measured by standardized tests such as the Modern Language Aptitude Test and Educational Testing Service Reading test, used the target language more often during study abroad and showed more gains in global oral proficiency, as measured by Oral Proficiency Interview scores, during study abroad abroad than those who had a lower level of grammatical and reading proficiency before SA.

Table 1. Factors affecting outcomes during study abroad

Factors that affect outcomes	Studies including each factor
Student-host family interactions	Knight & Schmidt-Rinehart, 2010; Martinsen, Baker, Dewey, Bown, & Johnson, 2010; Di Silvio, Donovan & Malone, 2014; Martinsen, Baker, Bown, & Johnson, 2011; Rivers, 1998; Schmidt-Rinehart & Knight, 2004; Segalowitz et al., 2004
Duration of study abroad	Davidson, 2010; Dwyer, 2004; Isabelli, 2004b; Isabelli & Nishida, 2005; Llanes & Muñoz, 2009; Sax, 2003; Serrano, Tragant, & Llanes, 2012
Out-of-class contact and social interaction with native speakers	Alvord & Christiansen, 2012; Dewey, Bown et al., 2013, 2014; Dewey, Ring et al., 2014; Fraser, 2002; Freed, 1990; Freed, Segalowitz, & Dewey, 2004; George, 2014; Linford, Zahler, & Whatley, 2013; Llanes, Tragant, & Serrano, 2012; Magnan & Back, 2007; Martinsen et al., 2013; Segalowitz et al., 2004
Use of the L1 while abroad	Stevens, 2011
Speaking the L2 with non-native speakers while studying abroad	Alvord & Christiansen, 2012; Magnan & Back, 2007; Martinsen et al., 2010
Non-interactive out of class contact such as television	Isabelli-Garcia, 2010; Segalowitz et al., 2004; Stevens, 2011
Motivation	Allen, 2010; Alvord & Christiansen, 2012; Hernandez, 2010; Llanes, Tragant, & Serrano, 2012; Martinsen, 2010; Martinsen, Alvord & Tanner, 2014; Trenchs-Parera & Juan-Garau, 2014
Age of the learner	Dewey, Bown et al., 2014; Llanes & Muñoz, 2013; Muñoz & Llanes, 2014
Sex of the learner	Brecht et al., 1995
Intercultural sensitivity	Martinsen, 2010; Martinsen & Alvord, 2012
Attitudes	Alvord & Christiansen, 2012; Llanes, Tragant, & Serrano, 2012
Cognitive processing and/or phonological memory	Freed et al., 2004; Lord, 2006; Marqués-Pascual, 2011; O'Brien, Segalowitz, Freed, & Collentine, 2007; Sunderman & Kroll, 2009
Formal instruction during study abroad	Cheng & Mojica-Díaz, 2006; Lord, 2010; Shively, 2011; Trenchs-Parera, 2009
Pre-study abroad language/grammatical proficiency	Brecht et al., 1995; Dekeyser, 2010; Dewey, Ring et al., 2014; Duperron & Overstreet, 2009; Freed, 1998; Magnan & Back, 2007; Segalowitz et al., 2004
Study abroad program type	Dewey, Bown et al., 2014; Dewey, Ring et al., 2013; Engle & Engle, 2003; Fraser, 2002; Martinsen et al., 2010

Dekeyser (2010) found through interviews and intensive observation that learners with insufficient proceduralized grammatical knowledge before study abroad were not able to automatize their grammatical knowledge during SA, making it nearly impossible to speak accurately and use even basic linguistic structures by the end of a six-week study abroad period. This implies that having a higher level of grammatical proficiency prior to studying abroad will lead to an increased ability to automatize grammatical knowledge and make more oral proficiency gains during SA. This is in accord with other research studies, such as Segalowitz et al. (2004) and Magnan and Back (2007), who argued that L2 learners' preparation for study abroad is crucial to predicting linguistic gains in the study abroad context because this context of learning is extremely complex, requiring learners to juggle many different types of communicative situations and pressures. Magnan and Back (2007), for example, found that the level prior coursework of L2 learners of French who studied abroad was a good predictor of performance in the study abroad context in that those who had more target language experience made the most gains during SA, supporting the hypothesis that their preparation prior to study abroad was crucial to their development in that learning context.

Other studies, though, have found that learners with lower levels of grammatical proficiency at the beginning of a study abroad program can sometimes achieve greater gains during study abroad than those with higher levels of proficiency (Freed, 1998; Regan, 2003). The argument for this type of finding is that advanced learners tend to perform closer to ceiling levels and have less room for improvement than lower-level learners. This, though, can depend both on what is being measured and what tasks are used to measure proficiency and linguistic competence, as certain task designs tend to lead to more ceiling effects than others (e.g., Freed, 1998; Mora, 2014). It is therefore important to continue to investigate whether grammatical proficiency is related to gains during study abroad because it may not be the same for each aspect of linguistic competence that is studied and for each methodological approach. With further research using different methods and studying a variety of aspects of linguistic

competence, we can begin to piece together a more complete understanding of how grammatical proficiency predicts L2 learners' outcomes during SA. Another important consideration is that it has yet well understood whether grammatical proficiency predicts performance in the study abroad context on measures of linguistic competence that are very different from grammar, such as phonology. This is important because it could help us understand better the connections between the acquisition of different linguistic competencies during study abroad and how they are similar and dissimilar. In other words, there could be a relationship between having a higher grammatical or lexical proficiency, for example, and higher performance on measures of speech perception, such as lexical decision, because lexical access is strongly related to the development of the L2 lexicon. Therefore, the current study will investigate whether language learners' grammatical proficiency, as measured by a grammar test, predicts speech perception patterns in the study abroad context both in terms of initial performance and change over time.

NS contact and Target Language use. The amount of contact that L2 learners have with NSs and language use are two of the extralinguistic factors that have been most frequently studied, most often via the Language Contact Profile (Freed et al., 2004) and similar questionnaires that are administered to L2 learners before and after a study abroad period. These types of questionnaires elicit information regarding the learners' language background prior to study abroad and the amount of time learners spend engaging in different language activities such as speaking the target language with native speakers, watching television in the L2, and reading in the L2. The hypothesis regarding these factors is that L2 learners who spend more time engaging with NSs of the target language in the target language will receive more authentic input and will therefore evidence greater linguistic gains over time during study abroad. In reality, though, studies have shown conflicting results for the effects of NS contact and target language use on gains observed during SA. Segalowitz et al. (2004) summarize the findings of their large-scale, multi-faceted research project that compared study abroad and at-home groups'

performance over the course of one semester on measures of vocabulary, pronunciation, grammar, communication strategies, oral proficiency and oral fluency according to extralinguistic factors measured by the Language Contact Profile. They found that the reported amount of activity in Spanish outside of class did not correlate significantly with oral fluency and oral proficiency measures for study abroad students. However, there were significant correlations between the use of communication strategies and the amount of time reported using the target language outside of the classroom and the amount of time reported speaking with host families in the target language. This showed that the more students communicated with NSs, the less likely they were to have gaps in communication and the need to use communicative strategies to fill those gaps. Additionally, study abroad students developed superior narrative discourse abilities compared to at-home students. Increased target language use during SA, though, did not significantly affect learners' performance on measures of discrete features of the target language, such as grammar and pronunciation. Segalowitz et al. (2004) argued that second language acquisition in different contexts of learning is much more complex than is often assumed, and that the *nature* of different communicative interactions in the target language and the diversity of input, rather than the overall time spent communicating in the target language, may be of more importance. This is because different communicative situations provide different types of input and feedback to the learner, which, in turn, could affect acquisition in different ways. For example, if a learner spends a lot of time alone but reports frequent interactions with shopkeepers in the target language, outcomes may be different than for someone who speaks the target language frequently with a NS who is a close friend even if the overall time reported using the target language is similar. This is because the NS friend will likely interact differently with the L2 learner than a shopkeeper and the diversity of input each learner receives will most likely differ.

Magnan and Back (2007) used a very similar questionnaire in order to predict linguistic gains according to contact with NSs and use of the target language by L2 learners. They found that learners'

living situation during study abroad (e.g., homestay, dormitory etc.) and interaction with media in the target language did not predict more gains, and that the only factor that did predict gains was prior coursework in French, indicating that language experience prior to studying abroad was more important than language use factors. This is in accord with Segalowitz et al. (2004) and Dekeyser (2010) in that they suggest that the target language experience and level of readiness that L2 learners bring with them into the study abroad context are crucial. Interestingly, Magnan and Back (2007) also found that speaking the target language (French) with Americans actually impeded linguistic gains during SA. Based on their results they suggested the L2 learners who study abroad be advised to take advanced coursework prior to study abroad in order to prepare well, and then avoid spending time with L1 peers and speaking the target language with them during SA. This study shows the importance of having a good measure of extralinguistic factors that may affect linguistic gains during study abroad because common assumptions about the study abroad context, such as the assumed benefit of living with a host family, do not hold true in every case (Rivers, 1998; Segalowitz & Freed, 2004; Wilkinson, 1996, 1998). Individual differences abound, even with learners in the same learning context. For example, two learners living with the same host family may use that opportunity differently depending on other factors such as age and language proficiency, which has been found by other researchers (Dewey, Bown et al., 2014).

Dewey, Bown et al. (2014) analyzed the factors that predicted L2 learners' NS contact and target language use during study abroad in six different study abroad programs. The tasks included a language log and the Study Abroad Social Interaction Questionnaire that they developed, among other instruments. They found that the learners' age, initial language proficiency level, and their development of social networks during study abroad all contributed to varying degrees to how much contact with native speakers they had and how much they used the target language during SA. Specifically, learners who were of a lower proficiency level sought out more opportunities to speak with NSs while more

advanced learners made more use of receptive language activities (e.g., watching TV in the target language). Older learners used the target language more than younger learners, which they hypothesized may have been the result of greater maturity or different motivations than those of younger learners. In terms of social networks, they found that study abroad program interventions such as requiring two hours per day of interaction with native speakers outside of class or volunteer activities in the community led to better social network formation and L2 use. In this way, the type of study abroad program type in which learners were enrolled and the programming that was incorporated into the curriculum was a significant predictor of language use because of different requirements between programs that led to differing amounts of target language input, which may, in turn, lead to greater linguistic gains.

Similarly, Martinsen et al. (2010) found that learners in different types of immersion programs (i.e., traditional study abroad, service-oriented study abroad, and at-home foreign language housing at an American university) showed different amounts of use of the target language. For example, learners in a traditional study abroad program used the target language more in the classroom while learners in a service-oriented study abroad program used the target language the most overall when both receptive (e.g., media use) and interactive language activities were combined. However, in this study they found that the overall amount of time spent using the target language (Spanish) by learners in all three groups did not correlate with higher ratings by native Spanish speakers for pronunciation, grammar, fluency, vocabulary, and comprehensibility as elicited by parts of the Oral Proficiency Interview and Texas Oral Proficiency Test. Furthermore, they found that learners who spent more time speaking the target language with their roommates that spoke the same L1 on a daily basis received higher ratings from the NSs than those who spent more time speaking the target language with NSs. This is the opposite of what Magnan and Back (2007) found, that speaking the target language with other L2 learners inhibited gains in Oral Proficiency Interview scores (i.e., OPI). Martinsen et al. (2010) hypothesized that this surprising

result may have come from differences in the types of interactions that learners have with peers that share the same L1 and NSs of the target language. In other words, interactions with NSs may involve more listening than speaking on the part of the L2 learner due to an unequal fluency level between learners and NSs, while interactions between L2 learners are on more equal ground and may allow for more speaking practice in a comfortable environment, resulting in greater oral gains. This appears to be more evidence to support Segalowitz et al.'s (2004) assertion that overall contact time with NSs may not always be sufficient to account for the nature of communicative interactions in the study abroad context and their benefits for L2 learners. Contact with other L2 learners may also provide certain benefits. This may also differ depending on what aspect of linguistic competence is being measured and how it is measured.

Dewey, Ring, Gardner and Belnap (2013) studied the development of social networks by 71 learners of Arabic during study abroad in Jordan and Egypt. They found that L2 learners believed that the greatest benefit came from the use of the target language with people with whom they had a close relationship, as opposed to people outside of their normal social circle. Personality and time spent with native speakers were the two greatest contributing factors to the development of social networks. Speaking with those outside of the main social circle also had a positive supplementary effect, though it was not as strong. They did not, though, test the L2 learners on any sort of linguistic measure to determine what effect this had on gains. In addition, they found a similar result to Martinsen et al. (2010) in that study abroad program interventions played a significant role in the development of social networks. For example, one program required two hours of speaking with locals per day. The two program locations also provided different opportunities for social network formation due to the locations of the students. In Jordan, they were near a university campus and their social networks consisted of mainly college-aged peers. In Egypt, on the other hand, the students lived near shops and a sports club and their social networks included local shopkeepers and staff members of the club. In this

way, the study abroad programs provided different opportunities and environments for social network creation, and thus different types of target language input, which may impact language gains.

George (2014), too, found that social networks played an important role, but she also investigated how this affected linguistic gains, specifically for the production of two dialect-specific phonological variants in Spain during SA. In this study, social networks played both a positive and negative role, depending on the individual. George (2014) found that those who had a stronger social network of Castilian Spanish speakers at the end of the period abroad used the Castilian /θ/ more than those without as strong of a social network with Castilian Speakers. On the other hand, half of the participants who never produced either of the dialect-specific variants reported having a social network that included at least one Puerto Rican (i.e., non-Castilian) Spanish speaker. The Puerto Rican students lived in the same dormitory as the L2 learners. Thus, the dialect of the native speakers with whom the learners had contact played a role in their use of Castilian variants.

In summary, the literature to date on extralinguistic factors in different contexts of learning has provided a conflicting account of the effects of these factors on language acquisition during SA. Studies have clearly shown that L2 learners differ widely in their development of social networks, target language use, and preparation for study abroad prior to arriving in the target language community. These issues are more complex than often assumed and it is clearly necessary to account for differences between individual learners, the different types of people with whom they have contact, and their target language experience prior to SA. However, one issue with some of the research to date on extralinguistic factors in the study abroad context is that, while studies such as Dewey, Bown et al. (2014), Martinsen et al. (2010) and Magnan and Back (2007) have investigated what factors predict the amount of language use and contact with native speakers a L2 learner will have during study abroad either without measuring linguistic outcomes or measuring general proficiency such as OPI scores or native speaker ratings. While this is certainly important, as it provides information regarding what

factors do lead learners to use the target language, this does not help us to understand how studying abroad, and the use of opportunities that come with that, may affect different specific aspects of linguistic (and sociolinguistic) competence in different ways. While Segalowitz et al. (2004) and George (2014) did study the effects of extralinguistic factors on linguistic gains, showing conflicting results, there are other research studies that report specifically on the types of linguistic outcomes that have been observed when comparing learners in different contexts of learning while also considering factors that predict outcomes. The next section will review the findings of study abroad research regarding linguistic outcomes for different aspects of linguistic competence and will present a gap in our current understanding of SLA during study abroad that the current study begins to fill.

Study Abroad Outcomes

Having discussed the importance of input, differences between contexts of learning related to input, and extralinguistic factors that can affect access to it, it is now necessary to review the development of research on context of learning over the years and what the general outcomes have been in order to understand where the field is going and how the current study contributes to our understanding of context of learning in SLA. Early studies on the acquisition of a second language in the study abroad context focused on global oral and grammatical proficiency (i.e., fluency, pronunciation, communication strategies, sociolinguistic skills, literacy skills), as well as students' beliefs and perceptions about study abroad and how those beliefs affected their learning (see Freed, 1998 for a review). The results of these early studies generally showed that studying a second language in the study abroad context had a significant positive influence on general linguistic skills, as shown by increases in Oral Proficiency Interview (OPI) scores for learners of Spanish, French and Portuguese (Foltz, 1991; Liskin-Gasparro, 1984; Magnan, 1986; O'Connor, 1988; Veguez, 1984). However, these studies

used global measures of oral proficiency such as the OPI rather than testing specific linguistic structures, and they often lacked at-home comparison groups.

Since then, more sophisticated methodologies have been used to study acquisition in the study abroad context, looking at more specific areas of linguistic and sociolinguistic competencies and comparing learning contexts (i.e., study abroad vs. AH). One important and consistent finding is that oral fluency, which is defined as rate of speech, length of utterances, time between utterances, compensation strategies, and pauses and hesitations, has frequently showed more gains by learners in the study abroad context than those in formal instruction contexts and is considered to be one of the primary aspects of linguistic competence to see gains during study abroad (Du, 2013; Freed, Segalowitz, & Dewey, 2004; Garcia-Amaya, 2009, 2012; Möhle, 1984; Mora & Valls-Ferrer, 2012; Perez-Vidal & Juan-Garau, 2011; Perez-Vidal & Sanz, 2014; Raupach, 1984; Segalowitz & Freed, 2004; Serrano et al., 2012; Valls-Ferrer, 2011; Valls-Ferrer & Mora, 2014). For instance, Valls-Ferrer and Mora (2014) collected fluency data from Spanish and Catalan speaking learners of English who were tested four times over the course of 2.5 years, including six months of formal instruction followed by a three-month study abroad period in the target language community, and ending with another six months of formal instruction. They found that the learners showed significant gains in oral fluency during the course of the three-month study abroad period, but that the formal instruction periods before and after study abroad did not have a significant effect on fluency scores. Importantly, though, gains were not equal for all participants, leading Valls-Ferrer and Mora (2014) to investigate individual factors such as initial fluency level and NS contact variables. They found that those with higher initial fluency scores also had higher scores than those with lower initial scores at all data collection times, but that learners in both groups made significant gains during the study abroad period. In other words, the fact that gains during study abroad were larger than gains during the formal instruction periods was not a result of the different initial fluency levels, but rather a result of the differences between the learning contexts (p. 133). They

also found that initial fluency level and the amount of contact with NSs during study abroad were positively related. The amount of contact a learner had with native speakers and their fluency gains were also positively related, though the effect was too weak to make a strong claim about this relationship.

Along similar lines, other studies have found that the learners in the study abroad context tend to show gains over time for listening comprehension (Beattie, Valls-Ferrer, & Pérez-Vidal, 2014), the acquisition of sociolinguistic competence (Barron, 2006; Kennedy, 2012; Regan, Howard, & Lemée, 2009; Thomas, 2004), lexical complexity/vocabulary (Ife, Vives Boix & Meara, 2000; Juan-Garau, Salazar-Noguera & Prieto-Arranz, 2014; Pizziconi, 2013; Serrano, Llanes, & Tragant, 2011; Serrano, Tragant, & Llanes, 2012), narrative abilities (Collentine, 2004; Duperron & Overstreet, 2009; Lafford, 2004), communication and language strategies (Duperron & Overstreet, 2009; Lafford, 1995, 2004), pragmatic competence (Bataller, 2010; Cohen & Shively, 2007; Felix-Brasdefer & Hasler-Barker, 2015; Hassall, 2013; Kinginger, 2008; Kinginger & Farrell, 2004; Schauer, 2007; Shively, 2011, 2013), intercultural learning and cultural sensitivity (Martinsen & Alvord, 2012; VandeBerg, Connor-Linton & Paige 2009) and sometimes writing abilities (Godfrey, Treacy, & Tarone, 2014; Pérez-Vidal & Barquin, 2014; Pérez-Vidal & Juan-Garau, 2009; Sasaki, 2007, 2009, 2011). What is common among most of these areas of linguistic competence for which significant gains during study abroad have been found is that most are broad aspects of communicative competence that indicate that study abroad learners tend to focus on being able to communicate meaning rather than acquiring specific morphosyntactic or phonological forms. It is not a coincidence, then, that studies that have investigated the acquisition of specific aspects of the interlanguage grammar (e.g., morphosyntax/phonology) have many times, though not in every case, found little to no significant positive effect of being in the study abroad context or even a more positive affect for being in the at-home context than the study abroad context.

With a small number of exceptions (e.g., Isabelli & Nishida, 2005; Juan-Garau et al., 2014; Marques-Pascual, 2011), studies have found that aspects of grammar that are the focus of classroom instruction (Cheng & Mojica-Diaz, 2006; Collentine, 2004; DeKeyser, 1991; Isabelli-Garcia, 2010; Kinginger, 2008), reading comprehension (Dewey, 2004), and sometimes writing fluency and/or general writing abilities (Freed, So, & Lazar, 2003; Lord, 2009; Sasaki, 2004) tend to lack significant gains during study abroad and even evidence more gains among learners in the at-home context in some cases. One hypothesis is that which was argued by DeKeyser (2010), that study abroad learners who do not make grammatical gains may lack sufficient grammatical knowledge before studying abroad, which could make it difficult to automatize their knowledge during study abroad (Brecht et al., 1995; DeKeyser, 2010; Lafford & Collentine, 2006). Also, the hypothesis mentioned above that at-home learners, by nature of learning primarily in a classroom context, pay more attention to form than meaning while study abroad students pay attention to meaning more than form due to a greater need for broader communication skills like fluency and communication strategies when living in the target language community, has been supported much of the research to date, including some studies on the acquisition of phonology during study abroad (e.g., Avello & Lara, 2014; Mora, 2014). Since the current study investigates the acquisition of phonology during study abroad, the literature on the acquisition of phonology in the study abroad context will now be reviewed.

Research on the acquisition of L2 phonological forms in the study abroad context has shown similar findings to the studies on the acquisition of morphosyntax that is emphasized in the formal classroom. In other words, results have been somewhat inconclusive, occasionally finding a significant effect for phonological acquisition in the study abroad context (e.g., Alvord & Christiansen, 2012; Martinsen & Alvord, 2012) but many times there is not, or there is only a significant effect when some other factor is taken into account. For example, Díaz-Campos (2004) was one of the first studies to investigate the production of specific L2 phonemes and allophones during SA, comparing study abroad

learners' and at-home learners' productions of Spanish word-initial stops (i.e., [p t k]), intervocalic fricatives (i.e., [β ð ɣ]), word-final laterals (i.e., dark or light [l]), and palatal nasals (i.e., [ɲ]). The participants were 26 study abroad students and 20 at-home students that read a paragraph containing 60 target words that included the target segments both before and after one academic semester in their respective contexts of learning. Importantly, he also administered the Language Contact Profile (Freed et al., 2004) to test the effects of language background, contact with native speakers, and language use factors on pronunciation gains during the semester. Overall, Díaz-Campos (2004) found that both the at-home and study abroad groups made gains over the course of the semester and that there was not a clear advantage for the study abroad context for pronunciation. He showed that the pronunciation of multiple Spanish phones during study abroad varied according by target phone and speech style. Greater gains were attested for voiceless initial stops and word-final laterals than intervocalic fricatives and the palatal nasal, showing that not all phonological processes were acquired at the same rate, even in the study abroad context.

In a follow up study on the effect of speech style with the same group of learners, learners in the study abroad context showed more native-like pronunciation in informal speech than the at-home learners, but the groups were not significantly different in more formal speech (Díaz-Campos, 2006). He hypothesized that this resulted from more opportunities to speak informally in the study abroad context than the at-home context. Other factors that predicted gains were having a higher number of years of formal instruction (i.e., 7+ years) being female, having higher OPI scores at the beginning and end of the semester, and having begun formal instruction in elementary school as opposed to junior high (the least native-like), high school and university levels (borderline). This last factor's significance indicates a significant effect for the age of onset of L2 learning, which has been found in other studies to be important for the acquisition of phonology and affecting degree of foreign accent (Flege, 1991; Flege, Bohn & Jang, 1997; Flege, Munro & MacKay, 1995; Flege, Yeni-Komshian & Liu, 1999). Díaz-Campos

(2006) is an important study in the literature due to its design that included a detailed questionnaire that was able to test the effects of different individual factors on pronunciation over time in the study abroad context.

In search of other factors that predict the acquisition of phonology during SA, Lord (2010) sought to determine whether it was immersion, explicit phonetics instruction, or a combination of the two that was best for acquiring Spanish spirantization during overseas immersion (i.e., the production of the phonemes /bdg/ as spirants [βðɣ] in intervocalic position and following liquids consonants). She found that immersion in the study abroad context alone did not predict more gains for the study abroad group than a group that only received explicit phonetics instruction about spirantization, but that when a third group of learners received explicit phonetics instruction of spirantization *combined* with being in the study abroad context, they made the most gains. In fact, other research, too, has shown that explicit phonetics instruction is beneficial for L2 learners' pronunciation, even outside of an immersion context (Aliaga-Garcia & Mora, 2009; Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999; Isabelli-Garcia, 2010; Pisoni & Lively, 1995).

Alvord and Christiansen (2012) also studied the acquisition of the pronunciation of Spanish intervocalic spirantization of the phonemes /b d g/ ([β ð ɣ]) by learners who spent two years abroad as missionaries. Thus, it was not a study about change over time but rather measured outcomes after the missionaries returned home. They administered two different production tasks: a read-aloud story followed by answering questions regarding the content of the story, for which they were instructed to ignore pronunciation, and reading a word list using their best pronunciation. They also filled out a questionnaire that inquired about their language background and included two surveys of cultural sensitivity and motivational intensity. The productions were analyzed spectrographically. They found that most of the learners attained a target-like pronunciation (81%), which was higher than what was found in previous studies on the acquisition of Spanish spirantization. Contrary to Díaz-Campos' (2006)

findings, they did not find an effect for style except for the production of /b/, which was more accurate in the story task (i.e., more casual speech). Other factors that predicted more accurate pronunciation were having five or more years of music instruction, having less years of previous Spanish instruction, spending less time speaking Spanish with English-speaking companions (i.e., like Magnan and Back, 2007), and having higher levels of cultural integration and motivational intensity. Their results suggest that even learners without much explicit instruction can acquire spirantization and that these extralinguistic factors exerted a strong influence on their accuracy of pronunciation. They also, contrary to Lord (2010), argued that formal instruction may have hindered pronunciation accuracy and that receiving poor Spanish input via English speakers impeded acquisition. An important difference, though, between this and previous studies is the length of time abroad. Two years is much longer than one semester and may have contributed to the high accuracy in pronunciation.

Martinsen and Alvord (2012) conducted a similar study in terms of investigating the effect of cultural sensitivity on L2 learners' pronunciation after a six-week study abroad program in Argentina. However, this study did not measure pronunciation acoustically, but rather relied on native speaker ratings of the learners' oral responses to two contextualized tasks taken from the OPI and Texas Oral Proficiency Test. While they did see a slight increase in pronunciation scores over time, the difference was not statistically significant. Regarding the cultural sensitivity factors, they found that pronunciation gains were positively correlated with the total cultural sensitivity score that the learners received at the beginning of study abroad and was highly correlated with a subsection of the cultural sensitivity questionnaire that represented learners' attitudes toward others. They concluded that positive attitudes toward other cultures predicted improvements in pronunciation.

Martinsen, Alvord and Tanner (2014) sought to determine how L2 learners' pronunciation at different levels of language instruction compared to NSs, including a group of learners who had a two-year experience abroad as missionaries. They also compared classroom learners' pronunciation in Oral

Proficiency Interviews to those who had been abroad for two years in a Spanish-speaking country. The learners were tested using the OPI and native Spanish speakers rated their degree of foreign accent on a sliding scale of zero to 100. They found that the pronunciation of students with the two-year experience abroad was much more similar to that of NSs than the pronunciation of students without such an experience. The learners who had been abroad, though, still exhibited significantly more foreign accent than the NSs. Interestingly, they found that the highest rated learner, a female student, did not have a two year experience abroad, but did study abroad for one semester. Additionally, this student reported that six months prior to SA, she created a plan to improve her pronunciation by listening to at least 15 minutes of Spanish newscasts or other media three to four days of the week and imitating the speakers' words and phrases, followed by reading aloud from texts while attempting to sound like the speakers to whom she had just listened. Thus, their results indicate overall that spending time abroad can help in the reduction of foreign accent, but that it is also possible to do so in other ways. It is important, also, to point out that the participants in their study spent a very long time abroad compared to the average study abroad student. Therefore, it is difficult to generalize their results to the traditional study abroad population. Also, a limitation of their study is that the use of NS ratings does not provide information as to why certain learners were rated higher or lower than others and what specific aspects of the L2 phonology develop during time abroad. Native speaker ratings are also limited in that they can often be representative of fluency factors more than actual pronunciation. It simply lends support to what Segalowitz et al. (2004) say about learners that study abroad, that they sound better, or more native-like, because they have developed superior narrative discourse abilities and have learned to hold conversations in the target language and have less need to fill gaps in communication (p. 13).

Avello and Lara (2014) is a recent study about the acquisition of English phonology during study abroad by Catalan and Spanish speakers who were learning English. The goal of Avello and Lara (2014) was to investigate segmental production and examine the role of another factor, length of stay in the

target language environment (i.e., three versus six months). The study focused on the quality and duration of the English vowel pairs /i:- ɪ/ and /æ - ʌ/, as well as voice onset times (VOT) for the voiceless stops /t/ and /k/ as produced by L1 Catalan and Spanish speakers from Barcelona who were studying abroad in England. Data was collected prior to studying abroad and after their return from SA. Prior to studying abroad all participants in the three-month group (N=25) had completed a formal instruction period, while the six-month group (N=8) had not. The speech samples were taken from a read-aloud task. Vowel tokens were extracted and measured for duration and quality. They also measured the voice onset time produced for the voiceless stops /t/ and /k/ in word-initial stressed position. The results showed that VOT and vowel quality and duration were not significantly affected during the study abroad period.

Avello and Lara (2014) concluded that, like most other studies of L2 phonological development during SA, their study did not show significant gains for the production of L2 phonemes. They support the hypothesis of other researchers about communicative nature of the study abroad context, where L2 learners may realize that accented L2 speech does not always hinder communication, leading them to focus on other communication skills that help them with fluent communication in the L2 and may reduce foreign accent (Martinsen, Alvord & Tanner, 2014; Segalowitz et al., 2004). In fact, Mora (2014), who studied the development of the perception of several English vowel and consonant contrasts during a study abroad experience, was led to the same conclusion by his results.

The L2 participants in Mora's (2014) study were also native Catalan and Spanish bilinguals who were tested on the perception of English vowel and consonant voicing contrasts using an AXB discrimination task at four time intervals (i.e., initial, after a period of formal instruction prior to study abroad, after the study abroad period, and a delayed posttest). 160 participants completed the AXB task at Time 1 (beginning of formal instruction), 102 at Time 2 (end of formal instruction, beginning of SA), 152 at Time 3 (end of SA), and 43 at Time 4 (delayed post-test). 66 participants completed the

perception tasks at Times 1, 2 and 3 and only 27 did so at all four data collection times. The analysis, therefore, compared the results between Times 1 and 2 (during formal instruction), between Times 2 and 3 (study abroad period), and for the 27 participants who completed all parts of the study, the delayed posttest at Time 4. Mora (2014) also included an English NS group (N=9) to serve as a baseline. Results showed high discrimination scores at all data collection times, but that some of the scores were different depending on the phonological contrast in question. The longitudinal data showed that discrimination scores were higher for vowel contrasts than consonant voicing contrasts overall, but that the vowel and consonant scores were highly correlated over time. Interestingly, there was a sharp increase in discrimination accuracy for both vowels and consonants after the formal instruction period that then remained steady through the study abroad period and the delayed posttest. Mora (2014) found that the learners with a low initial level of discrimination ability made the most gains during the formal instruction period and those who made gains during the study abroad period had a high initial level at Time 1 and did not make significant gains during the formal instruction period. In sum, the learners' gains during the formal instruction and study abroad periods were dependent on their initial level at Time 1 and there was a ceiling of around 85 percent that Mora hypothesized that no one could likely surpass without explicit instruction and practice that directs their attention to acoustic cues. The combination of immersion and instruction is something that other researchers have pointed to as having the strongest effect on gains (Lord, 2010).

To summarize, studies on the acquisition of phonology and pronunciation in immersion contexts indicate that some gains are possible, but that they have mostly been observed among those who spend a long time in the target language culture (e.g., missionaries), or gains are only observed when certain other factors like speech style, formal pronunciation instruction, and learners' target language background are taken into account. In other words, there is no all-encompassing benefit for phonology and pronunciation due to being immersed in the target language culture. Mora (2014), based on the

results of his speech perception study, suggested that perhaps the typical three-month study abroad period was not sufficient time to see perceptual gains or that, like Avello and Lara's (2014) hypothesis, the learners focused more on broader communicative functions of language rather than segmental perception. In fact, positive findings from the same large-scale research program for fluency (Valls-Ferrer & Mora, 2014) and listening comprehension (Beattie, Valls-Ferrer, & Pérez-Vidal, 2014), when considered alongside other findings showing little to no significant results for the production (Avello & Lara, 2014) and perception (Mora, 2014) of phonological contrasts during the study abroad period support this hypothesis and the findings of Segalowitz et al. (2004) that the broad aspects of linguistic competence are what tend to show the most gains over time in the study abroad context.

Another important hypothesis, though, that Avello and Lara (2014) put forth regarding the underwhelming results of phonological and pronunciation studies in the study abroad context is that the phonological units that they and other researchers have studied may not have been salient enough to be noticed by the learners if they really are focused on meaning in the study abroad context. Unfortunately, research to date has lacked speech perception studies on the acquisition of phonology in the study abroad context. Mora (2014) is one of the very few studies to have investigated speech perception during immersion in the target language environment. Without a perception component, it is impossible to know whether the phonological units under study are noticed or not by L2 learners. For this reason, the focus on production that has permeated research on L2 phonological acquisition in the study abroad context has been missing half of the picture. Perhaps in a few or even many of the studies to date, production has been impeded by a lack of perception.

Finally, another issue with the current state of research on context of learning is that most of the studies to date have focused on the acquisition of aspects of the L2 grammar and phonology for which there is little to no variation among native speakers. Only recently is this beginning to change. Avello and Lara's (2014) suggestion that some phonological units that have been studied in the

literature may not have not been salient enough to be noticed by learners may apply more to linguistic units that do not carry social meaning than to those that do (i.e., sociolinguistically variable phonological forms). We know that native speakers exhibit a wide range of variation according to linguistic (e.g., grammatical/phonological context) and extralinguistic (e.g., place of origin, age, sex, socioeconomic class) factors (Labov, 1972), and it follows that those L2 learners who study abroad will encounter variation of some type. Since L2 learners have been shown to focus more on being able to communicate successfully with native speakers of the language in the study abroad context than the minutiae of grammar or phonology that do not vary among NSs, it may be that a case that a variant that is specific to the target language community in which a learner studies, for example, could be more salient to them due to not having been exposed to it before and being a variant that distinguishes some dialects and people from others. L2 learners may see at least the comprehension of dialect-specific variants as more necessary for communication with native speakers with whom they come in contact than certain other linguistic forms that are not variable. This is what is suggested by Geeslin (2011b), who, in her review of the literature on the acquisition of variable linguistic forms in L2 Spanish, remarked that learners who have not acquired sociolinguistic variation, “may well be limited in their expressive ability, unable to correctly interpret language directed to him or her or, worse, he or she may project an inappropriate social image in certain contexts” (p. 462). In other words, not acquiring sociolinguistic variation, either with regard to comprehending the variation (i.e., perception) or producing the variation, may impact the learners’ social relations with others. For this reason it is important to determine if the acquisition of linguistic forms that are variable during study abroad is different than acquiring linguistic forms that are categorical, because we may be missing half of the story when it comes to the acquisition of a second language in different learning contexts.

Therefore, the current study seeks to contribute to our understanding of the relationship between the acquisition of sociolinguistic variation and context of learning in SLA. To accomplish this,

the current study investigates factors that influence the perception of a dialect-specific phonological variant, /s/-aspiration in Western Andalusian Spanish, by L2 learners of Spanish who studied abroad for one semester in Seville, Spain, where the aspiration of /s/ is widespread (Villena Ponsoda, 2008). L2 learners, who were enrolled in two different study abroad programs in Seville, were compared to an at-home group of university students in the U.S., who did not have exposure to Western Andalusian /s/-aspiration before or during the semester. Western Andalusian /s/-aspiration was chosen as the target variable because it is prevalent in the target city of Seville, it is unique among /s/-aspirating dialects in its phonetic articulation, and the perception of this variety of /s/-aspiration by learners of Spanish has not been previously studied. In Western Andalusian Spanish, the duration of aspiration of coda /s/ preceding a voiceless stop tends to be short or non-existent and the quality varies between breathy voicing and voiceless glottal frication. Additionally, the VOT of a voiceless stop following a coda /s/ tends to be longer than normal voiceless stops in Spanish and the stop closure of the voiceless stop can be different for these stops than for stops that are not preceded by coda /s/. (O'Neill, 2010; Ruch, 2013; Ruch & Harrington, 2014; Torreira, 2006, 2007, 2012).

In order to understand the motivations behind the research questions and methodological design of the current study, it is necessary to first review the literature on the acquisition of variable linguistic forms in a second language and what we know from a small body of studies to date about how variable forms are acquired in a study abroad context.

Chapter 2 – Review of the Literature

This chapter will review the previous literature that motivates the investigation into the second language perception of a dialect-specific phonological variant in a study abroad context. The first section will review the literature on the acquisition of sociolinguistic variation in a second language, including a general overview of variation in linguistic systems, the acquisition of variable linguistic structures in Spanish as a second language, the acquisition of variation in immersion contexts, and the perception of L2 dialect-specific phonological variants. This will be followed by a description of the sociolinguistic variation of the phoneme /s/ throughout the Spanish-speaking world, as well as a phonetic and sociolinguistic description of the target variant of the current study, /s/-aspiration in Andalusian Spanish specifically. Then, theories of L2 phonology that are relevant to the current study will be reviewed and hypotheses will be made regarding how L2 learners are expected to perceive Andalusian /s/-aspiration based on its acoustic properties. This final section will also include a discussion of the two speech perception task paradigms that make up part of the methodology of the current study. The chapter will conclude with the presentation of the research questions that guide the current study.

The Acquisition of Variable Linguistic Forms in a Second Language

Variation in Linguistic Systems

Research in the field of sociolinguistics over the years has described in detail the variation that is inherent to linguistic systems and has investigated many factors that condition variation. Variation occurs where multiple linguistic forms can be used to fulfill a single grammatical function or a single grammatical form fulfills multiple functions in the language. In these variable contexts, native speakers of any given language will employ certain forms over others depending on various factors, such as the linguistic context (e.g., phonological context, stress, grammatical category of the word) and the social context (Labov, 1972). Social context is defined by factors such as the formality of the speech context,

the characteristics of the interlocutor with whom the speaker interacts and the characteristics of the speaker (i.e., age, sex, socioeconomic class, education level). While many studies have focused on the variation that occurs in first language linguistic systems, another body of research has investigated whether second language linguistic systems also exhibit variation and what factors predict the use of variable forms in a second language, finding that L2 grammars do exhibit variation and that it is conditioned by many of the same types of factors that condition L1 variation.

Early second language acquisition research viewed the errors that learners made as nuisances that should be corrected immediately and they were seen as either the result of poor instruction or the improper development or reinforcement of habits from a behaviorist perspective (Skinner, 1957). However, Corder (1967) argued that learners' errors are in fact meaningful in that they provide evidence of the development of the second language linguistic system at a particular point in time. Learner "interlanguage," as Selinker (1972) called it, is a systematic linguistic system that changes over time and shows variation at any given point in time. Selinker (1972) argued that the interlanguage can be observed when a learner is focused on communicating meaning (i.e., in conversation) rather than when the learner is focusing on linguistic form (i.e., classroom drills). Following this logic, early studies on the interlanguage system mimicked sociolinguistic research methods at the time by manipulating task type and therefore speech style in order to study interlanguage variation according to the amount of attention paid to speech (Dickerson, 1975; Tarone, 1985; Tarone & Parish, 1988). For example, Dickerson (1975) tested the production of the English /z/ by 10 Japanese speakers who were learning English in the U.S. over a nine-month period using a free speaking task (i.e., least careful/formal speech), a dialogue reading task and a word list reading task (most careful/formal speech). She found that the learners' productions of /z/ varied depending on the speech task, as well as the phonological environment, which are both factors that have been shown to condition L1 variation.

As sociolinguistic research methods have become more sophisticated, so have methods for studying L2 variation. More factors have been added to statistical models in order to better explain patterns of variation in both first and second language systems. For example, Adamson and Regan (1991) studied the production of the English variant [iŋ] versus [in], as in *walkin’/walking* by Cambodian and Vietnamese learners of English as a second language. Their findings showed that the L2 learner’s sex, the degree of monitoring of their speech, and the grammatical category of the item that contained the target variant all significantly affected the learners’ production of [iŋ] versus [in]. Beebe (1980) showed that another factor, the characteristics of the interlocutor with whom an L2 learner communicates, plays a role in the production of phonological and other types of linguistic forms. This is a very similar finding to what has been reported in studies of interlocutor effects on L1 variation (e.g., Rickford & McNair-Knox, 1994), again bolstering the evidence that variation in L1 and L2 linguistic systems are conditioned in the similar ways.

The examples described above are cases of ‘Type II variation’ or ‘horizontal variation’ (Rehner, 2002), which, for L2 learners, is the variation between “aspects of the target language where native speakers display sociolinguistic variation, that is, they alternate between variants as a function of linguistic and extra-linguistic factors” (Mougeon, Rehner & Nadasdi 2004:409). In other words, Type II variation is a part of both L1 and L2 grammars. This is in contrast to Type I variation, which characterizes the developmental stages through which L2 learners progress over time as they acquire the second language and restructure the linguistic system (Ellis, 1985).

The L2 acquisition of Type II variation is part of *sociolinguistic competence* and it has been shown to be a crucial aspect of L2 competence (Bayley, 2005; Bayley & Regan, 2004; Bayley & Tarone, 2012b; Canale & Swain, 1980; Geeslin, 2011a, 2011b; Geeslin & Gudmestad, 2011; Mougeon, Nadasdi & Rehner, 2010; Mougeon, Rehner & Nadasdi, 2004; Regan, Howard & Lemée, 2009; Rehner, Mougeon & Nadasdi, 2003). This is because learners who have not acquired sociolinguistic variation, “may well be

limited in their expressive ability, unable to correctly interpret language directed to him or her or, worse, he or she may project an inappropriate social image in certain contexts” (Geeslin, 2011b: 462). Geeslin (2011b) makes a crucial point about two important aspects of sociolinguistic competence. The first aspect of sociolinguistic competence mentioned by Geeslin (2011b) could be labeled ‘production.’ If a L2 learner does not *produce* certain sociolinguistic variants, his or her expressive ability could be limited. In other words, L2 learners may be limited in the range of meaning that they can communicate in the L2 if they do not produce certain variants that native speakers produce. The other important aspect of sociolinguistic competence mentioned by Geeslin (2011b) could be labeled ‘perception.’ In the words of Geeslin (2011b), a learner who lacks sociolinguistic competence might be, “...unable to correctly interpret language directed to him or her...” (p. 462). L2 learners who lack sociolinguistic competence may misinterpret the basic meaning associated with a linguistic form (e.g., if it is a Type II variant among native speakers with whom the learner has contact), the intended meaning of the interlocutor or the social values associated with certain linguistic forms.

Taking as an example the target variant of the current study, /s/-aspiration, L2 learners, upon receiving input that contains /s/-aspiration, must acquire the linguistic meaning associated with the aspirated variant (i.e., that it is an allophone of the phoneme /s/), the phonological information about /s/-aspiration (i.e., phonotactic constraints), and the social component of /s/-aspiration (i.e., an understanding of the social values associated with /s/-aspiration). On the production side, if the learner incorporates /s/-aspiration into production, we must take into account both the linguistic (i.e., phonetics/phonology) and the sociolinguistic (i.e., whether they use /s/-aspiration appropriately) aspects of learners’ production.

The use of the terms *production* and *perception* here to describe sociolinguistic competence is purposeful, as it draws a parallel to the terms *production* and *perception* in the study of L2 phonology. In L2 phonology, the *production* of an L2 phone refers to whether a learner uses a certain L2 phone, how it

is articulated (i.e., phonetics), and the learner's acquisition of the phonological constraints on its use. *Perception* in L2 phonology, on the other hand, refers to how L2 learners categorize L2 phones in the L2 phonological system and process the meaning of words from the phonemes and allophones upon which they are built. In the study of L2 phonological variation, then, it is important to make a distinction between the *production* and *perception* of a phonological variant and to define the goals of the study. When considering the acquisition of variable forms in a second language, perception, then, has two components: the perception of linguistic meaning that is associated with a variant (i.e., categorizing it correctly) and the perception of the social and contextual constraints surrounding the variant. The current study is concerned with the perception of a dialect-specific phonological variant, /s/-aspiration, specifically the aspect of perception that is related to the *linguistic* meaning associated with /s/-aspiration. In other words, it is not within the scope of the current study to investigate L2 learners' perception of the social and contextual constraints related to /s/-aspiration, but rather how the acoustic properties of /s/-aspiration are categorized in the L2 phonological system and how /s/-aspiration affects L2 learners' processing of words. In this way, the current study investigates *one part* of the sociolinguistic competence of L2 learners, but there is clearly more that can be investigated in future research.

Like research on context of learning, most studies to date on the acquisition of variable grammatical or phonological forms in a second language have focused on production. While this is clearly an important aspect of sociolinguistic competence, few studies have focused on the perceptual aspect of acquiring variation, both in the sense of the perception of linguistic meaning and the social and contextual constraints attached to variable forms. The next section will review the research on the acquisition of variable forms in L2 Spanish and the motivation for conducting a perception study related to the acquisition of a dialect-specific phonological variant in a study abroad context.

The Acquisition of Variable Structures in Second Language Spanish

Before reviewing the recent and growing body of research on the acquisition of sociolinguistic variation during study abroad specifically, it is important to review the range of variable forms in L2 Spanish that has been researched as well as general findings in order to understand considerations regarding research design for studying the acquisition of variable L2 forms, what learners have been capable of in terms of acquiring variable forms, what types of factors are at work in the acquisition of variable forms, and where this dissertation fits in the current L2 variation research landscape.

Many studies throughout the history of the field of SLA have focused on the acquisition of linguistic forms in obligatory contexts that require the use of one form for one linguistic function (e.g., grammatical gender). Analyzing this type of data requires a comparison of the form produced by an L2 learner and the form that is required by the grammar of the target language in a given context (i.e., correct vs. incorrect). However, as has already been discussed, not all grammatical and phonological contexts are obligatory. Some allow for multiple forms to be used for one function or one form to be used for multiple functions, and native speakers' form selection is determined by a variety of factors. Therefore, a correct versus incorrect dichotomy is not sufficient to capture L2 acquisition of a variable form. In response to the need for new methods of analysis to account for variable contexts, Geeslin (2000) implemented a method of analysis that accounted for the variable context by not analyzing learners' productions as correct or incorrect but, instead, analyzed both the rates of use (i.e., proportion of use of different forms) and the factors that predicted the use or preference of a certain variable form by L2 learners in comparison to the rates of use and factors that predicted the use of a variable form among native speakers of the target language (see Geeslin, 2011b for a review of studies of this type). In this type of analysis, predictors of use can be both linguistic (e.g., grammatical, phonological) and extralinguistic factors (e.g., age, years of L2 study, age of onset of L2 learning, motivation, contact with

NSs, target language use) that condition the preference for one form over the other in a given variable context.

With these methodological considerations in mind, research on the acquisition of variable structures in Spanish has expanded in recent years to include studies on mood selection (Geeslin & Gudmestad, 2008a, 2010; Gudmestad, 2006, 2008, 2010a, 2010b, 2012a, 2012b, 2014; Kanwit & Geeslin, 2014), subject expression (Geeslin & Gudmestad, 2008a, 2011; Geeslin, Linford, Fafulas, Long, & Diaz-Campos, 2013; Gudmestad & Geeslin, 2010; Linford, 2009, 2014; Linford & Shin, 2013), verbal objects (Geeslin et al., 2010; Malovrh, 2006, 2008), future time expression (Gudmestad & Geeslin, 2011, 2013; Kanwit & Solon, 2013; Solon & Kanwit, 2014), present progressive (Fafulas, 2010, 2012, 2013; Geeslin & Fafulas, 2012), present perfect versus preterit for past time reference (Geeslin et al., 2012; Geeslin, Fafulas, & Kanwit, 2013; Whatley, 2013), and more recently phonological variation (Geeslin & Gudmestad, 2008b; George, 2014; Knouse, 2012; Ringer-Hilfinger, 2012; Schmidt, 2009, 2011) and pragmatic variation (Félix-Brasdefer & Koike, 2012; Félix-Brasdefer & Hasler-Barker, 2015).

In general, research to date has found that learners can often reach or at least move in the direction of native-like rates of use or preference of a variable form as they progress in their grammatical proficiency (Geeslin, 2000; Geeslin and Gudmestad, 2008a; Gudmestad, 2012b; Killam, 2011), or due to exposure to variable forms over time in the input through immersion (Geeslin et al., 2010, 2012; Geeslin, Fafulas, & Kanwit, 2013; Kanwit & Solon, 2013; Linford, Zahler, & Whatley, 2013; Salgado-Robles, 2011, 2014; Salgado-Robles & Enrique Ibarra, 2012; Willis, Geeslin, & Henriksen, 2009). At times the path to acquisition is not linear, but the learners still may eventually arrive at NS norms or at least approximate them. For example, Gudmestad (2012b) showed that learners can progress, and then regress before eventually reaching native-like use of the subjunctive. However, it is not always the case that learners approximate or reach native-like norms. In fact, multiple studies have found the

opposite to be true (Geeslin & Gudmestad, 2008a; George, 2014; Gudmestad & Geeslin, 2011; Geeslin, Fafulas & Kanwit, 2013; Knouse, 2012; Ringer-Hilfinger, 2012; Whatley, 2013).

These studies have found that even advanced learners do not always produce the same rates of use as NSs or move away from NS norms over time for both morphosyntactic variants and the production of phonological variants (Geeslin and Gudmestad, 2008a for the subjunctive; George, 2014; Gudmestad & Geeslin, 2011; Geeslin, Fafulas, & Kanwit, 2013 for only one group; Knouse, 2012; Ringer-Hilfinger, 2012; Whatley, 2013). There are multiple potential explanations for this, one of which is that studies sometimes only follow learners for a short period of time that may not be long enough to see the restructuring of the linguistic system. Another reason that has been reported is that, like native speakers, some learners have been shown to make a choice to not use a certain variant when the individuals that make up their social network do not use that variant (e.g., George, 2014; Ringer-Hilfinger, 2012). This shows that even when learners receive input containing variation they do not necessarily begin to use the variant.

In addition to findings that have shown that learners sometimes do not arrive at NS norms, studies have shown the same learners can be native-like in the use or preference of one type of variable form while being non-nativelike for another, and that different types of research tasks can provide different information about acquisition (e.g., Geeslin & Gudmestad, 2008a). For this reason it is important to study different kinds of variable structures using a variety of task types in order to better understand the process of acquiring variation in a second language as a whole.

To summarize the findings of L2 variation research to date, we have learned that in order to study the acquisition of variable structures it is necessary to account not only for how often learners produce, prefer, or perceive a certain variant, but also to account for factors that predict their use, preference, or perception. These factors can be linguistic or extralinguistic in nature. The study of linguistic factors helps us to understand whether L2 learners' use, preference, or perception of a variant

is determined by the same or different factors compared to native speakers. This is important because learners may use, prefer, or perceive a variant with a similar frequency to native speakers without completely understanding what drives NS use or preference of a variant, or they may not use or prefer a variant with a similar frequency to NSs (i.e., production), but may be acquiring the constraints that guide native speakers' form selection (i.e., perception). Extralinguistic factors are imperative because they help researchers predict how much input learners receive and/or what they do with input that they do receive. Through the study of linguistic and extralinguistic factors and their relationship to the acquisition of variation in a second language, we have been able to learn that when learners receive input that includes sociolinguistic variants they are capable of acquiring these variants at least to some degree, and that target language experience is important because acquisition of variable forms in the L2 generally occurs at more advanced levels of proficiency. This is at least partially attributable to learners' access to input that contains these variants, which tends to be limited in a traditional classroom setting. Without input there can be no acquisition. For this reason, it is important to study those learners who are better positioned to have access to input that contains variable forms. One way of doing so is to investigate the acquisition of L2 variation in immersion settings.

When we study L2 learners' acquisition of variation in the L2 in immersion settings, then, there are a few important considerations with regards to input. As discussed in Chapter 1, the study abroad context does not guarantee access to input or an equal access to input for all L2 learners (e.g., Diao, Freed & Smith, 2011). For this reason, the study of extralinguistic factors that can lead to more or less access to input, as well as to differentiate learners' who might use opportunities for input in different ways (e.g., Dewey et al., 2014), is essential. An additional input-related consideration for immersion contexts is the nature of the sociolinguistic variant that is the focus of the research. The preference and use of certain variants among native speakers can be determined solely on the basis of geography, or it can be determined by a combination of geography and social and contextual factors such as age, sex,

speech style, and socioeconomic class. When social and contextual factors are not at work in the use of a variant by NSs and only geography determines its use, a L2 learner will be exposed to input when in contact with any NS who is from that region. An example of such a variant is the Castilian Spanish phoneme /θ/. However, with phonological variation that is governed by both regional and sociolinguistic factors, the content of the input depends both on where the learner is, the individuals with whom the learner has contact, and in what types of speech contexts learners have contact with NSs. When investigating the acquisition of the latter type of variant, for which sociolinguistic factors also play a role in its use by NSs, it is even more essential to account for differences between individual learners and the types of contact that they have with native speakers.

With these considerations in mind, we will now review a body of recent Spanish and French studies that have used a variety of methods to investigate how L2 learners who study in different target language regions and are exposed to different input with relation to sociolinguistic variants acquire the variants, and what types of factors predict acquisition. First, findings related to morphosyntactic variants will be discussed, followed by findings for the production of phonological variants, and ending with the perception of phonological variants. In this way, the following review more specifically targets those studies that are most similar in purpose to the current study and will explain how the current study seeks to answer questions that remain unanswered.

The Acquisition of Sociolinguistic Variants in Immersion Contexts

To date, research on the acquisition of sociolinguistic variants in immersion contexts has focused primarily on morphosyntactic variation in Spanish and French. Research on Spanish structures has investigated the acquisition of variable direct object pronoun expression (Geeslin, García-Amaya, Hasler-Barker, Henriksen, & Killam, 2010; Salgado-Robles, 2011, 2014; Salgado-Robles & Enrique Ibarra, 2012), future-time expression (Kanwit & Solon, 2013), perfective past time reference (Geeslin, Fafulas, & Kanwit, 2013; Geeslin, García-Amaya, Hasler-Barker, Henriksen, & Killam, 2012), past-tense expression

(Whatley, 2013), and subject pronoun expression (Isabelli, 2004a; Linford, Zahler, & Whatley, 2013; Lopez-Ortega, 2003). Researchers of French have investigated the acquisition of the deletion of the negative particle *ne* (Dewaele, 2004a; Regan, 1995, 1996, 1998; Regan, Howard, & Lemée, 2009; Rehner & Mougeon, 1999; Thomas, 2004), the variable use of the subject pronouns *nous* and *on* (Rehner, Mougeon, & Nadasdi, 2003; Sax, 2003), subject-verb agreement (Howard, 2006), pronouns of address (Dewaele, 2004b; Lyster & Rebuffot, 2002), and a combination of these and other variables (Howard, 2012) during study abroad.

Research on the acquisition of morphosyntactic sociolinguistic variation in French during immersion in the target language community has led to important findings using sociolinguistic methods such as Labovian-style sociolinguistic interviews and role plays. First, and most importantly, learners of French have been shown to make significant gains for the production of morphosyntactic sociolinguistic variants over time in the study abroad context (Regan, 1996; Howard, 2012), after having a prior study abroad experience (Sax, 2003), and as a result of other types of prolonged contact with NSs of the target dialect (Dewaele, 2004a; Rehner & Mougeon, 1999). Though the L2 learners did not quite arrive at the NS norms for the production of morphosyntactic variants such as the deletion of the negative particle *ne*, the use of *nous* versus *on*, and variable expressions of futurity, we have seen that they have moved in the direction of the NS norms, particularly in the vernacular speech style (Regan, 1996; Sax, 2003). For example, Sax (2003) found that those who spent the most time abroad were the closest to NS norms and that the effect of speech style was the strongest for those who had been abroad the longest. Apart from speech style, these studies have found other factors condition the variation among learners, such as social factors (Rehner & Mougeon, 1999), the duration of study abroad (Sax, 2003), number of years of French study prior to the university level (Sax, 2003), and speaker-interlocutor dynamics such as age differential, sex, and whether the speaker and interlocutor are both L2 speakers or one L1 speaker and one L2 speaker (Dewaele, 2004a).

Studies on the acquisition of variable morphosyntactic structures in Spanish during immersion in the target language environment have shown similar trends to the study of variation in French, though the data collection methods have been quite different. Specifically, the studies of learners of Spanish have used written contextualized tasks rather than oral interview or role play data (with the exception of Salgado-Robles, 2011 and Salgado-Robles and Enrique Ibarra, 2012). Like the French studies, though, L2 learners have moved toward native-like norms for variable structures over time (Geeslin et al., 2010, 2012; Geeslin, Fafulas & Kanwit, 2013; Salgado-Robles, 2011, 2014; Salgado-Robles & Enrique Ibarra, 2012), sometimes having to decrease their use or preference for the variable structure to approximate NSs (Geeslin et al., 2010, 2012) and other times having to increase it (Geeslin, Fafulas & Kanwit, 2013; Kanwit & Solon, 2013).

The factors that have been shown to predict L2 patterns for variable morphosyntactic structures in the study abroad context include linguistic factors (e.g., *telicity* or *animacy of the subject* for the selection of *leísmo*: Geeslin et al., 2010; *time of action* for the use of the present perfect or preterit for past time reference: Geeslin et al., 2012), and individual factors such as previous study abroad experience, improvement over time on a grammar test (Geeslin et al., 2012), years of Spanish study (Geeslin et al., 2012) and exposure to different patterns of variation in different dialect regions (Geeslin, Fafulas & Kanwit, 2013; Kanwit & Solon, 2013; Salgado-Robles, 2011, 2014; Salgado-Robles & Enrique Ibarra, 2012). In other words, L2 learners with a prior study abroad experience, more improvement over time on a grammar test, a higher final grammar test score and more years of Spanish study have been found to evidence the most acquisition of target sociolinguistic variants in Spanish over time. In addition to these factors, we have learned that the location of one's study abroad program matters because of the differences in the input conditions associated with native speakers' use of a certain variant in different places. Salgado-Robles (2011), for instance, found that L2 learners who studied in Seville, Spain and Valladolid, Spain did not differ in their use of the dative object pronoun *le* and the accusative *lo* at

the beginning of a semester abroad, but did differ significantly at the end of the semester in their rates of use of these pronouns as they converged toward the norms for how NSs in each region used them. This type of finding has also been reported by Geeslin, Fafulas and Kanwit (2013) and Kanwit & Solon (2013) for learners of Spanish studying in Spain and Mexico.

There have also been recent studies on the acquisition of phonological variants in immersion contexts in both French and Spanish that will be described in more detail given that the current study focuses on the acquisition of a phonological variant rather than a morphosyntactic one. Beginning with studies on the production of phonological variants by L2 learners in the study abroad context, Thomas (2004) investigated the production of four morpho-phonemic features of French by learners who stayed at home in Canada and learners who studied in France. She distinguished between one that is universal in French (i.e., compulsory liaison) and one that is variable (i.e., schwa). A semi-directed oral test that required a description of a cover from *The New Yorker* magazine, a description of the student's university for advertising purposes, and a past tense narration of the student's autobiography, as well as two read-aloud texts showed that both groups were similar for compulsory liaison, but only those who studied in France began to follow the sociolinguistic norms of native speakers, particularly younger native speakers, in that they reduced their use of schwa and reduced their use of optional liaison. The at-home group did just the opposite and increased the use of schwa in every category, indicating that the at-home learners followed the norms as taught in the classroom.

Howard, Lemée, & Regan (2006) studied the acquisition of /l/ deletion in French and the effect of contact with native speakers. 15 L2 learners that had studied for one academic year in France and four that had never gone to France participated in sociolinguistic interviews that elicited 2,214 tokens of /l/. Howard, Lemée, and Regan found that learners who had spent one year in France deleted /l/ significantly more (33%) than at-home learners (6%), indicating that their exposure to /l/ deletion in France led to more native-like production of this sociolinguistic variant. However, they did not reach the

native speaker rate of /l/ deletion, which was 88 percent. Furthermore, though the mean NNS rate of deletion were much lower than NS rate, the hierarchy of constraints that governed the deletion of /l/ was similar between the NS and NNS groups, indicating that the learners who went to France began to acquire sociolinguistic constraints on the deletion of /l/.

The first study to investigate the L2 acquisition of phonological variation in Spanish was Geeslin and Gudmestad (2008b), who studied the frequency of use of the Castilian theta [θ] and /s/-aspiration among stateside learners of Spanish at the university level in order to determine how many learners produced these variants, in what contexts they produced them, the role that L2 proficiency played, and the role that a previous study abroad experience played. 130 university-level learners of Spanish from five different course enrollment levels completed a language background questionnaire, an 11-item grammatical proficiency test, and a monologic role play with six scenarios presented using PowerPoint.

The results showed that very few learners produced [θ] and /s/-aspiration during the role play task. Only nine out of the 130 learners produced theta and only five produced /s/ aspiration, varying from one use to 41 uses by a given participant. Additionally, learners in the two lowest enrollment levels never produced either target phone. The third level was where the production of [θ] was first observed and the fourth level was where the production of /s/-aspiration began. They also showed that a prior study abroad (SA) experience (including length of time abroad and recentness) and a desire to assimilate to the target culture did not play a significant role in the production of the variants. A study abroad experience was not necessary for a learner to use the variants and also did not necessarily lead directly to their use. This was shown because, even though all learners who had a study abroad experience and produced [θ] and /s/-aspiration had studied in Spain, there were also seven learners who had studied in Spain that did not use the variants. Their findings also did not indicate that a minimum length of stay was required in order to produce dialect-specific phonological variants. It is notable that Geeslin and Gudmestad (2008a) found that even exposure to sociolinguistic variation does not automatically lead to

use of a variant. Learners not only have to learn *how* to produce dialectal variants (i.e., phonetically), but also the phonotactic constraints on their use and the social values associated with them. They then must make a decision to produce the phonological variants, which can be affected by various individual factors such as motivation or attitudes (e.g., Ringer-Hilfinger, 2012). They concluded that, like variable morphosyntactic structures, the production of variable phonological processes are acquired late in the acquisition process, if at all.

Henriksen, Geeslin and Willis (2010) studied the acquisition of regional intonation patterns of five L2 learners of Spanish during a seven-week immersion program in Spain according to different utterance types. Data collection occurred via a background questionnaire, a grammatical proficiency test, a vocabulary test, and a computerized intonation elicitation task that was presented via PowerPoint using pictures, animations, dialogue balloons, and words. There was also a contextualized sentence reading task in which they read sentences that were preceded by a discourse context and were designed to prompt the production of declarative, absolute interrogative, or pronominal interrogative sentences. They found that that some learners modified their intonation contours toward the norms for the NSs from León over the duration of the program, but that there was significant variability among the five learners. Some learners modified their tonal inventory by introducing new strategies over time, resulting in more variability, while others did not produce as many new strategies but were more consistent. They note that this finding fits nicely in to what is known about SLA in general, that as the L2 grammar beings to include new features, variability increases until the more target-like feature supplants the previously-used feature. Importantly, their findings are connected to research on the acquisition of variation in SLA because intonation varies among native speakers, requiring learners to acquire native-like constraints on use, which usually leads to later acquisition, as has been shown for morphosyntactic structures.

A few recent small-scale studies have investigated the production of the Castilian theta /θ/ by learners during a study abroad program compared to at-home groups. The voiceless interdental fricative phoneme /θ/ is produced in most Spanish-speaking regions of the Iberian Peninsula, but is not produced in Latin American Spanish. While these studies suffer from methodological issues that significantly limit their generalizability, namely small sample sizes, they do confirm the findings of other research that variable phones are rarely produced by learners, even after exposure. Additionally, they provide qualitative data that lends some insight into why learners do or do not produce regional phonological variants.

Ringer-Hilfinger (2012) collected data using a cross-sectional design that included six learners who were tested four months prior to study abroad in Spain, three learners who were tested two months into a four-month study abroad program, and four learners who were tested six months after SA. She sought to determine how study abroad affected learners' awareness, attitudes, and use of /θ/, as well as the role of awareness on their attitude toward NS use of /θ/ and how their attitudes toward NS use of /θ/ affected their own use. The participants completed a background questionnaire, a matched guise task, a read-aloud text, a semi-directed informal interview, and a self-assessment of language use outside of the classroom. Overall, Ringer-Hilfinger found very low production of [θ] (2.9%, N=6/209) despite half of the participants reporting frequent /θ/ use, indicating that either the learners misrepresented their frequency of use of /θ/ or the tasks and/or research environment were not suitable for discovering it. The study abroad learners exhibited more awareness of /θ/ than non-study abroad learners. However, increased awareness did not lead to more use. The study abroad experience did not significantly affect their attitudes toward NSs or NNSs who produce /θ/. However, they did form opinions regarding their own /θ/ use. They reported that peer pressure and linguistic insecurity contributed to their nonuse of /θ/. Other factors that influenced the use of /θ/ were increased contact

with NSs who did not use /θ/ and having an instructor that did not use /θ/, indicating that the social networks of the learners played an important role in their use or non-use of /θ/.

Knouse (2012) is another recent study that investigated the production of /θ/ by L2 learners at home and abroad. The 15 study abroad learners were studying for six-weeks in Salamanca, Spain, where /θ/ is used, and they received /θ/ in the daily input from instructors. The 10 at-home learners were taking a six-week Introduction to Hispanic Linguistics summer course in the United States and were all minoring in Spanish. Four of the 10 at-home learners had previously studied abroad, three in Santander, Spain, and one in Seville, Spain. Importantly, the instructor of the at-home class spoke with a Castilian accent, produced /θ/ in 92 of 111 possible contexts that observed by the researcher during class, and also explicitly instructed the class on the use of /θ/ in Spain. Thus, both groups received input, though the study abroad group also received additional /θ/ input outside of the classroom. There were two speaking tasks that the participants completed: reading a short newspaper article and responding to open-ended questions. They then completed the Pronunciation Attitude Inventory (Elliot, 1995). Knouse analyzed the productions auditorily and, when necessary to differentiate between [s] and [z] productions when voicing could not be discerned, used a spectrogram. There was a high interrater agreement between Knouse and a native speaker.

Overall, Knouse found a very low rate of use of /θ/ by study abroad learners even with a large number of total possible phonological environments (1.7%, N=36/2,119), and she found no use of /θ/ by at-home learners between the beginning and the end of a six-week period. Knouse argued that one probable reason for the low overall use of /θ/ is that it is not necessary for communication, at least in most cases. Perhaps the learners were too consumed by other linguistic and cultural challenges to be able to incorporate a regional phoneme into their speech. However, half of the study abroad group was able to produce /θ/ at least once, indicating that something was occurring in their linguistic system that had perhaps not had time to be incorporated categorically into their phonological inventory. There was

a statistically significant difference between the two groups given that no at-home learners produced /θ/ despite having it explained to them and having an instructor that used /θ/. Knouse concluded that the study abroad context contributed to what little production of [θ] there was by the study abroad group, but that other factors impeded the further incorporation of /θ/ into the learners' linguistic system. These were factors such as task type, graphemic representation, grammar proficiency level, lexical stress, studying Spanish as a major/minor, housing situation while abroad, and a pronunciation attitude inventory score. She found that studying Spanish as a major or minor and, contrary to other studies like Geeslin and Gudmestad (2008), being an intermediate or beginner learner led to more use of /θ/ than the advanced learners. She proposed that perhaps the advanced students' phonological inventories were less accommodating to new phonemic categories due to having studied the language for a longer period of time than the beginner or intermediate learners, or that they were focusing on more sophisticated aspects of communicative competence.

George (2014) is the most recent study on L2 Spanish phonological acquisition during SA. She studied the production of /θ/ and the uvular fricative /χ/ by learners studying in central Spain for one semester. Both of these phonemes are region-specific, as they are part of the phonological inventories of central and northern Spain. She collected data at the beginning, middle, and end of the semester from 25 learners who studied abroad in Toledo, Spain using a guided conversation with a Castilian Spanish speaker, a read-aloud text, a list of words that was read aloud, a language background questionnaire, the Language Contact Profile (Freed et al., 2004), a social networks questionnaire, and a motivation and attitudes questionnaire. She found five students who used /θ/ frequently, eight who did not use /θ/ at all, and 12 with very infrequent use of /θ/. Their use of /θ/ did not change significantly over the course of the semester. Even the high-frequency users did not use /θ/ as much as NSs, who used /θ/ 100 percent of the time in possible contexts. The learners used /θ/ more in the word list task

than the reading passage or conversation tasks, which presented the least use of /θ/. She also found overgeneralization of /θ/ to contexts in which it is not normally used.

For the other phoneme, /χ/, eight learners used it frequently and 17 infrequently. The low-frequency user group did not change over time, while the high-frequency user group approached significance in its change over time, primarily from Time 1 to Time 2 (mid-semester). As was the case for /θ/, their use did not reach native-like levels. Task type was a significant factor at Time 1 and Time 3, and the learners produced /χ/ most in the word list task, less frequently in the reading passage task, and least frequently in the spontaneous conversation task. This is a similar finding to Regan (1996) in that the regional variant was most commonly produced by learners in the most monitored speech style rather than the informal, unmonitored style. As Regan (1996) argued, this indicates that learners recognized /χ/ as a target prestige form in this region of Spain and were careful to produce it when monitoring. However, they had not yet restructured their phonological systems enough to be able to produce it in unmonitored, informal speech.

George (2014) also analyzed the effects of extralinguistic factors on the use of these two phones. Learners with a higher proficiency level produced more /χ/ than those with a lower level, but the same was not true for /θ/. Some students lived in a dormitory with Puerto Rican students and claimed that one or more of the Puerto Rican students were in their social network. These students did not use the target phones and were likely accommodating to the Puerto Rican speakers more than the Castilian dialect. Some students reported a desire to speak like Castilian NSs, but were in the low frequency group. George (2014) hypothesized that they may have needed explicit instruction along with immersion to improve their use of the target phones. This could be true given that having had a previous Spanish instructor from Spain or having traveled to Spain previously correlated significantly in a positive way with /θ/ production at all three data collection times.

To review, studies on the acquisition of morphosyntactic variation in the study abroad context have found that, though they sometimes do not reach native-like norms within the time span of data collection, L2 learners often move in the direction of NSs over time for the rates of use of a given variant and in terms of the factors that constrain the use of a variant (or preference based on written contextualized tasks). This is important because it shows that learners can become sensitive to NS norms through exposure and that they can begin to restructure the interlanguage to comprehend and incorporate these variants. We have also seen that the content of the input that learners receive is extremely important, as native speakers' rates of use of a given variant and sociolinguistic factors will differ depending on where a learner studies (e.g., Salgado-Robles, 2011). Nevertheless, where a learner studies is not the only important factor. We know this because research on the production of phonological variants during study abroad has found that learners often stray away from producing them. In fact, learners behave a lot like native speakers in that they do not produce a phonological variant when their social networks consist of people that do not produce it (e.g., George, 2014; Ringer-Hilfinger, 2012), showing that learners accommodate most to those who make up their social network. This means that extralinguistic factors are very important for the acquisition of variable forms in a second language and that we cannot assume that simply going to a certain location will result in the acquisition of a variant. But an important question remains about whether the perception and production of sociolinguistic variants by L2 learners are affected differently by extralinguistic factors when learners have access to variable input. Clearly, the production of phonological variants involves a conscious choice to use the variant once a learner begins acquiring the meaning of a variant and the constraints on its use (both linguistic and social). Even though L2 learners may not use a certain variant, they may at least be acquiring sociolinguistic competence in the sense of correctly interpreting a variant's linguistic and social meanings. We now turn to recent studies that have investigated the perception of dialect-specific phonological variants to determine how they are processed by L2 learners.

The Perception of L2 Dialect-specific Variants

Of the very few studies to date on the perception of dialect-specific phonological variants in a L2, most have investigated the acquisition of languages other than Spanish. For example, Escudero and Boersma (2004) studied how L1 Spanish-speaking learners with exposure to two different dialects of English (i.e., Scottish and Southern British) and native speakers of English from Scotland and Southern England perceived the English vowels /i/ and /I/. In these two dialects there is a significant difference in F1 height for the vowel /I/, as well as durational differences. The test vowels were synthesized to create a continuum based on F1 values and duration. In general, the learners' perception patterned according to the dialect to which they had been exposed. More specifically, Escudero and Boersma found a difference between the two learner groups in terms of the acoustic cues (i.e., spectral information vs. duration) that each used to discriminate the vowels. Scottish L2 listeners preferred spectral cues while the Southern British English L2 listeners preferred both spectral and durational cues. They also found a difference between beginner learners and more advanced learners in that the beginner learners preferred spectral cues and the advanced learners preferred durational cues.

Baker and Smith (2010) studied the perception and production of the French vowels /i y u/ in two different dialects, Quebecois and European French, by L1 speakers of English. These vowels in the two dialects differ both in terms of the general acoustics and a specific acoustic cue (assibilation) that is particular to Quebec French. The learners of French spent two months studying in an intensive program in the United States, which was followed by a stay of 18 to 22 months in Quebec, Canada (N=10) or Toulouse, France (N=10). The first task was a discrimination task that tested the words /di/ (dit), /du/ (doux), and /dy/ (du) as produced by two female NSs from Quebec and two from southern France. The learners were instructed that they would hear German and French vowel tokens and had to discriminate between the vowels. The goal was to determine whether the learners could discriminate between the French vowels no matter the dialect. The results showed that the Quebec learners exhibited more

accuracy than the European French learners for the vowels produced by both the Quebec and European speakers. They concluded that the Quebec learner group made use of the additional cue from Quebecois French while the European French learners did not.

A second experiment tested whether the Quebec learners could extend their perception to contexts that do not allow for the use of the assibilation cue as well as whether both groups could perceive the vowels of other dialects also. The task was an identification task that required the listeners to hear a stimulus and choose using orthographic representations on the screen whether they heard the word *dit*, *doux*, or *du*. The results showed that the Quebec learners again were more accurate than the European learners for the Quebecois French /y/. This was in the absence of the assibilation cue, indicating that exposure to Quebecois French helped these learners to extend their perception to vowels outside of the assibilation context. Interestingly, the European learners were more accurate at identifying the vowels produced by the European speakers than the Quebec speakers, while the Quebec learners discriminated the vowels of both dialects equally well. The authors concluded that the assibilation cue in Quebecois French aided the perception of the learners exposed to that dialect overall, whereas the same effect was not seen for the European French learners. Finally, a production task showed that the learners' production of the vowels patterned according to the dialect to which they had been exposed, and again the Quebec learners were the most native-like in their production. Baker and Smith concluded that, "the dialect to which a learner is exposed may play a significant role in L2 acquisition" (p. 732).

The same authors conducted another study (Smith & Baker, 2011) to test the discrimination, identification, and production of German vowels of two different German dialects by L1 speakers of English who spent 16 months in Northern Germany or Southern Germany/Austria. The vowels were /i/ /y/ and /u/. English does not have the vowel /y/ and the two dialects of German differ in their production of these vowels. In Southern German the contrasts /i/-/y/ and /y/-/u/ are not as distinct

acoustically as they are in Northern German. In Southern German /y/ can also lose rounding, which creates a near-merger with /i/. The discrimination task was a same-different task in which the learners heard vowel pairs and had to determine whether two tokens, spoken in Northern and Austrian German, contained the same or different vowels. The identification task required the participants to select *die*, *dü*, or *du* on a computer screen depending on which vowel they think they heard. Finally, they also tested production. The results showed that the learners from both groups discriminated and identified the Northern German vowels most accurately, but that the Southern German group was even more accurate than the Northern German group for perception of the Northern vowels. The authors argued that perhaps exposure to more dialects of a language helps perception overall, even for dialects to which a learner is not directly exposed. Additionally, they found that the salience of the difference between /y/ and /u/ in Southern German did not help the learners in that group as much as hypothesized because the de-rounding of /y/ seemed to hinder their perception the most, indicating that the dialect to which a learner is exposed can have both positive and negative effects on their perception and acquisition of phonology.

Similarly, Fox and McGory (2007) unexpectedly found that exposure to a dialect over time did not significantly help the identification of Southern English vowels by a group of Japanese-speaking learners of English who lived in the United States for at least two years. The task was an identification task and ten English vowels were identified using the target words *hid*, *hayed*, *head*, *had*, *who'd*, *hood*, *hoad*, *hod/hawed*, and *hud*, which were displayed on a computer screen. Surprisingly, the group of Japanese-speaking learners of English living in Alabama (Southern dialect) perceived Northern (Ohio) vowels more accurately than the Southern vowels even after having lived in Alabama for two years. Likewise, comparing a group of Japanese speakers living in Ohio to the group living in Alabama, there was no significant difference between the two groups for the identification of Southern vowels. The Ohio L2 group also had slightly higher accuracy than the Alabama group for the Ohio vowels.

On a production task, the Ohio L2 group's production of the ten vowels was more similar to Ohio NSs than the Alabama group's productions were to Alabama NSs. Fox and McGory hypothesized that the positive results for the Ohio vowels and negative results for the Alabama vowels for both L2 groups may have resulted either from exposure to the media, which predominantly features vowels more similar to the Ohio vowels than the Southern vowels, or that the Ohio vowels were easier to identify due to a greater difference in duration for tense-lax vowel pairs. This is relevant because in Japanese, vowel length is phonemic and these listeners may have been more attuned to the durational differences in Ohio vowels. They also suggested that sociolinguistic factors such as the individuals with whom the learners had contact and their attitudes toward the southern variety of English may have contributed to the Southern English L2 group's lack of gains in the perception and production of Southern English vowels.

Schmidt (2009), on the other hand, found that L2 learners of Spanish were able to improve comprehension of speech that contained dialect-specific phonological variants from the Dominican Republic after only three weeks of exposure during a short-term study abroad program. She studied 11 English-speaking L2 learners of Spanish one week before and two days after returning from a three-week stay in the Dominican Republic, testing the comprehension of words and phrases that contained four variable phonological features that are common in Dominican Spanish: syllable- and word-final [s] deletion (e.g., *pastas* 'pastas' /'pas.tas/→['pa∅.ta∅]), lambdacism (e.g., *hablar* 'to speak' /a.'blar/→[a.'blal]), word-final nasal velarization (e.g., *pan* 'bread' /'pan/→['pan]), and word-final [d] deletion (e.g., *edad* 'age' /e.'dad/→[e.'da∅]). The participants heard words and phrases extracted from sociolinguistic interviews with native speakers of Dominican Spanish. A small subset of control words and phrases also came from interviews with male NSs from other regions of the Spanish-speaking world. After hearing a word or phrase the participants wrote down in Spanish the word or phrase that they had heard along with its meaning in English. The word task contained eight words with [s] deletion, four with

lambdacism, four with /n/-velarization, and two with [d] deletion. The phrase task included four with one or more words exhibiting [s] deletion, three with lamdacism, two with [d] deletion, and one with nasal velarization. Participants also completed a vocabulary familiarity task and language background and proficiency measures.

Schmidt (2009) found that by the end of the three weeks the learners were able to significantly increase the accuracy of their comprehension of Dominican Spanish, having become more familiar with the dialect-specific phonological forms. However, at Time 1 and Time 2 they still exhibited more accuracy with the control words and phrases than the Dominican words and phrases and they also made slight gains for the control words and phrases over time. More specifically, their comprehension varied depending on the target phone and task type. The deletion of final [s] and lambdacism were more difficult to comprehend than [d] deletion and /n/-velarization. Finally, the learners made more gains on the word-level comprehension task than the phrase-level comprehension task, likely because the phrase-level task often included more than one phonological feature in the same phrase. Schmidt (2009) concluded that dialect familiarity has a significant effect on listening comprehension by L2 learners, which had also been found in previous research (Tauroza & Luk, 1997). The difference between the findings of Schmidt (2009) and Fox and McGory (2007) could be a result of differences in L1-L2 phonological categories when comparing Japanese to English and Spanish to English.

Continuing to explore the perception of dialectal variation in L2 Spanish, Schmidt (2011) conducted a large-scale study that tested the perceptual identification of the aspirated variant of the Spanish phoneme /s/ in Venezuelan and Argentinian Spanish in word-internal coda position by L2 learners of Spanish (N=215) and native speakers (N=47). The two native speaker groups were from the non-/s/-weakening region of Columbia (N=27) and the /s/-weakening region of La Rioja, Argentina (N=20). The second language learners were L1 English-speakers who were enrolled in five different levels of university-level Spanish courses. The levels were first-semester beginner (Level 1, N=57),

second-semester intermediate (Level 2, N=52), third-year introductory topics courses on Spanish culture and literature (Level 3, N=59), fourth-year advanced topics courses in Spanish literature and linguistics (Level 4, N=26), and graduate students of Spanish literature (Level 5, N=21). All participants completed a forced-choice identification task in which they heard isolated target non-words containing [s] (i.e., sibilant) or [h] (i.e., the aspirated variant) in word-internal coda position (N=14) and required them to choose the correct orthographic representation for the word they heard from seven options on the computer screen, including 'none' (i.e., <baspe>, in which <s> represents the allophones [s] and [h] orthographically). The coda response options were <s>, <r>, <l>, <f>, <n>, or no coda (i.e., there was no coda consonant, as in <bape>). In addition, there were 50 control items that tested other coda conditions such as /l/, /r/, /m n/ (i.e., nasals), /f/, and no coda (i.e., there was no coda consonant in the stimulus). Schmidt (2011) also included 94 distracter stimuli testing other sound contrasts in different word positions. The stimuli were recorded by two native speakers of Spanish, one male from Caracas, Venezuela, and one female from Buenos Aires, Argentina. The participants also completed a background questionnaire in order to elicit information regarding their previous study abroad experience and overall target language background.

Schmidt (2011) found that the learners' identification of Spanish /s/-aspiration (i.e., identifying [h] in coda position as orthographic <s>) increased concomitant with an increase in course level, beginning with the third level and increasing in the fourth and fifth levels (i.e., late acquisition). Level 1 and Level 2 learners did not accept the aspirated variant as a valid variant of /s/ to any significant degree. Identification was also affected positively by prior exposure to dialects that exhibit /s/-aspiration as reported on the language background questionnaire. Those who had not previously been exposed to /s/-aspirating dialects through study abroad did not identify [h] as <s> to the same extent as those who had. She also found that those who had more target language experience in general tended to exhibit higher accuracy rates of identifying [h] as <s>. The results for the two native speaker groups

showed that the NSs from a /s/-aspirating region (Argentina) identified [h] as <s> significantly more than the /s/-maintaining speakers, as would be expected. However, it was not necessary for a native speaker to produce /s/-aspiration to be able to identify it correctly. A small number of NSs from the non-aspirating group exhibited more accuracy in identifying [h] as <s> than the rest of their group. Schmidt found that in these cases some of the native speakers had contact with other native speakers from /s/-aspirating regions and it appeared that it influenced their perception of the target variant.

Schmidt (2011) is relevant to the current study because it was the first large-scale study to investigate the L2 perception of dialect-specific allophonic variation in Spanish, showing that both experience with the L2 and exposure to specific varieties can lead to more accurate identification of the target variant. These findings are in accord with previous research on the acquisition of morphosyntactic variants in that acquisition started to occur as learners became more advanced and were exposed to input from native speakers within the target-language community. Interestingly, though, the results differ from those found in studies on the production of phonological variants by L2 learners. In other words, it appears that, while /s/-aspiration is not usually produced by even advanced learners (Geeslin & Gudmestad, 2008b), it can be perceived by them. Also, in Geeslin and Gudmestad's (2008b) study, a study abroad experience was not related to learners' production of dialect-specific phonological variants, but Schmidt (2011) found that a study abroad experience in a dialect region in which the target variant, /s/-aspiration, occurred was significantly related to the perception of /s/-aspiration. In this way, it seems that the production and perception of dialect-specific variants are indeed different and further research needs to continue to explore these differences by incorporating speech perception into the methodological repertoire of research on context of learning and the acquisition of phonological variation.

Schmidt (2011) has made a valuable contribution to the literature, yet there are two important points that must be addressed with further research in order to continue to investigate how the

perception of dialect-specific variants occurs. One reason further research is needed is that Schmidt's methodology, and that of other studies to date on the perception of dialect-specific variants, have focused on orthographic identification and have not measured how learners process real lexical items containing /s/-aspiration (i.e., lexical access). We know from previous research that different tasks can produce different results and also reveal different aspects of the acquisition of a given linguistic structure or form (e.g., Geeslin & Gudmestad, 2008a). In the case of the perception of dialect-specific variants, while it is clearly important to understand how learners associate the target phone with orthography in non-words, we must also seek to learn whether phonological variation presented in real lexical items hinders their processing of those items, which would have implications for learners' communication with native speakers in the target language community. Another necessary follow-up to Schmidt (2011) is to test the longitudinal development of the perception of a dialect-specific variant by learners who study abroad and are exposed to the target variant in the input. This is necessary because a longitudinal design also allows for the investigation of the effects of other extralinguistic factors that might affect learners' perception of dialectal variants over time.

In sum, the literature reviewed thus far has provided a rationale for conducting such a study on the perception of a dialect-specific phonological variant over time in a study abroad context by showing the scarcity of research on the perception of dialect-specific variants both generally and particularly in the study abroad context, where input is generally greater than in other contexts of learning. The review of the literature has also revealed limitations of the methods used to date, which have focused primarily on the identification of variants according to their orthographic representations and, with the exception of a few studies such as Sumner and Samuel (2005, 2009), have not tested the online processing of real words containing phonological variants. Therefore, the current study builds on the design and findings of Schmidt's (2011) study by taking a longitudinal approach to the study of the L2 perception of /s/-aspiration in Spanish, investigating factors that predict change over time for the same learners in a study

abroad context, and incorporating two tasks (i.e., identification and lexical decision) to measure two different aspects of the perception of /s/-aspiration.

It is now necessary to describe the sociolinguistic and phonetic nature of the variant that is the focus of the current study, /s/-aspiration in Andalusian Spanish, in order to be able to predict what kind of input learners will receive during study abroad and how they will perceive /s/-aspiration initially and then after a period of exposure. Accordingly, the next section reviews the literature on the phoneme /s/ in Spanish, the types of /s/-variation that are found throughout the Spanish-speaking world, the sociolinguistic factors that drive /s/-variation, and a phonetic and sociolinguistic description of /s/-aspiration in Andalusian Spanish.

Spanish /s/

The phoneme /s/ in Spanish has two primary voiceless sibilant allophones: apicoalveolar [s̺] and laminoalveolar [s], depending on dialect region (Mason, 1994). The apical [s̺] is most common in the northern and central varieties of Peninsular Spanish as well as a few regions of Colombia, while the laminal [s] is used in Southern Spain (i.e., Andalusia) and the great majority of Latin America (Mason, 1994, p. 18). The Spanish /s/ can also undergo voicing ([z]), typically before a voiced consonant within a word (*chisme* ‘gossip’ [tʃiz.me]), across word boundaries (*dos monos* ‘two monkeys’ [doz.mo.nos]) and before a word-initial glide (*las llamas* ‘the flames’ [laz.ja.mas]) (Hualde, 2005). However, recent research has shown just how variable /s/-voicing actually is. For example, Schmidt and Willis (2011) found that the duration of voicing was variable in Mexican Spanish, as /s/ was often only partially voiced and there was a significant amount of variation among individual speakers. Some speakers tended toward voicing most of the segment while others tended toward voicelessness. There was also a lack of voicing before a voiced consonant in 37 percent of the cases, and they found voicing in intervocalic position and preceding a voiceless consonant, where it is not expected to occur according to prescriptive accounts.

The voicing of /s/ has also been found to occur in word-final pre-vocalic position in Highland Ecuadorian Spanish (*dos aves* ‘two birds’ [doz.a.βes]) (Lipski, 1989; Strycharczuk, Van ‘T Veer, Bruil & Linke, 2014) and Highland Columbian Spanish (Garcia, 2013), though to a lesser degree than among Ecuadorian speakers. Interestingly, Garcia (2013) found that speech rate had a significant effect on the percentage of the sibilant that was voiced. Finally, Bradley (2006) and Sessarego (2011) report on the production of /sr/ clusters in multiple varieties of Spanish, showing that a coda /s/ preceding a trill /r/ in onset position of the following syllable frequently results in an assibilated trill rather than two distinct segments.

In addition to the variants of /s/ mentioned above, many dialects of Spanish evidence /s/-weakening. Evidence from historical linguistics suggests that Spanish /s/-weakening originated between the 16th and 19th centuries in Andalusia (Fontanella de Weinberg, 1990; Lipski, 1995; Mason, 1994). Spanish dialectologists began describing /s/-weakening using auditory judgments in the early 1900s (Navarro Tomás, 1918; Henríquez Ureña, 1921). These studies described the geographic distribution of /s/-weakening, finding that it tends to occur mostly in lowland and coastal areas of South and Central America, the Caribbean, the Canary Islands, southern Spain (i.e., Extremadura, Andalusia, Murcia) and parts of Mexico (Mason, 1994). The weakening of /s/ has not been reported for highland regions of Mexico, Central America, the Andean regions of South America and northern regions of Spain in general (Hualde, 2005). However, it has been reported in parts of Madrid, Toledo, La Mancha, Cuenca, Salamanca, Rioja Baja and Avila, Spain (Quilis, 1999). The weakening of /s/ has also been shown in many studies to vary according to language-internal factors (i.e., phonological environment, word position, prosody, lexical frequency) and social and stylistic factors such as age, sex, socioeconomic class and speech style within and among /s/-weakening regions. Given the frequency of /s/-weakening in Spanish and its vast regional and social variation, it was selected as the target variable for the current study.

Consonant weakening (i.e., lenition) is a very common process among the world’s languages (Lavoie, 2000). In Spanish, the aspiration of /s/ (e.g., *pasta* ‘paste’ [pah.ta]) and elision (e.g., [pa∅.ta])

are the most common variants of /s/-weakening (Cedergren, 1973; Lipski, 1984; Mason, 1994). However, aspiration and deletion are not the clean-cut categories the International Phonetic Alphabet makes them out to be. Aspiration has been shown to be characterized by a combination of various acoustic cues due to differing relations between glottal and oral gestures in different dialects. The most common aspirated variant is voiceless aspiration (*pasta* 'paste' [pa^h.ta]), in which the glottis remains open, the vocal folds do not vibrate, and there is very little constriction in the oral cavity, producing high airflow that is not as strident as for [s], which has a tighter and longer oral constriction (Mason, 1994). This will henceforth be referred to simply as 'aspiration.' Other variants include breathy voicing instead of voiceless aspiration (O'Neill, 2010), as well as the post-aspiration ([pa.t^ha]) and affrication ([pa.t^sa]) of a voiceless stop that tends to follow a coda /s/ in Andalusian Spanish (Moya Corral, 2007; O'Neill, 2010; Ruch, 2012, 2013; Torreira, 2006). In Andalusian Spanish, an aspirated /s/ can also be followed by a geminate stop (i.e., longer stop closure) ([pa^ht.ta]), particularly in Eastern Andalusian Spanish (Bishop, 2007; Gerfen, 2002).

The elision of [s], too, is not a clear-cut category, as its elision often leads to the use of compensatory mechanisms such as vowel lengthening to compensate for the loss of [s]. For example, Carlson (2012), Hammond (1978) and Figueroa (2000) found that lengthening the preceding vowel served as a cue for the loss of [s] in word-internal syllable-final position in Andalusian (Carlson), Miami-Cuban (Hammond) and Puerto Rican (Figueroa) Spanish. Vowel harmony has also been reported as a compensatory mechanism for the loss of [s] in Eastern Andalusian Spanish (Gerfen, 2002). Other recent studies have shown that the elision of [s] leads to lengthening the closure of a following stop in Puerto Rican Spanish and a change in manner of articulation of voiced stops (Luna, 2010; Galarza, Delgado-Díaz & Willis, 2014). Galarza, Delgado-Díaz & Willis (2014) studied the productions of eight Puerto Rican university students in the context of /s/ followed by the three voiceless and three voiced stops in Spanish across word boundaries using a picture description task. They found that when /s/ in this

context was produced as [s], elided, and in a small number of cases, aspirated. They found that the duration of all stops was systematically lengthened when [s] was elided. Stop closure lengthening for voiceless stops was the primary acoustic cue for an elided [s]. They also found that voiced stops were phonetically different following an elided [s] than in intervocalic contexts, both in terms of closure duration and manner of articulation (i.e., fortition rather than lenition). They suggest future perception studies to determine the effects of these mechanisms serve a functional purpose upon elision of [s].

Findings from studies of /s/-weakening in the Spanish-speaking world have shown that /s/ can be weakened in multiple different positions within a word, some of which are common while others are rarer (e.g., Brown & Torres Cacoullos, 2002, 2003; Jiménez Sabater, 1975; Lipski, 1984, 1999; Torreira & Ernestus, 2012). For example, /s/ is most commonly weakened in syllable-final position preceding a consonant either in word-internal (1) or word-final (2) position preceding a consonant in word-initial onset position of the following word. It can also be weakened before a pause (3), before a vowel in cases in which the /s/ is at the end of a word which precedes another word that begins with a vowel (4)¹, or in syllable-initial position (5), which is the least common environment for /s/-weakening.

- | | |
|--|--------------------------|
| 1. <i>¿Cómo está?</i> ('How are you?') | [ˈko.mo.eh.ˈta] |
| 2. <i>Hablas mucho</i> ('you speak a lot') | [ˈa.blah.ˈmu.tʃo] |
| 3. <i>Tengo chocolates</i> ('I have chocolates') | [ˈteŋ.go.tʃo.ko.ˈla.tɐh] |
| 4. <i>Los árboles</i> ('the trees') | [lo.ˈhar.βo.leh] |
| 5. <i>Cosa</i> ('thing') | [ˈko.ha] |

That /s/-weakening occurs most often in syllable-final position preceding a consonant is not surprising given findings for many of the world's languages, which tend to weaken consonants in pre-consonantal syllable-final position (Terrell, 1979: 608). This is for two primary reasons. First, /s/ has been shown to be aspirated most often in word-internal coda position while it is more often deleted in

¹ In this context it re-syllabifies to the onset of the following syllable.

word-final position (Cedergren, 1973; Hammond, 1980; Lafford, 1986). Second, the phoneme /s/ has been shown to be aspirated more often before consonants than before vowels or a pause (Alba, 2000; Cedergren, 1973; Lipski, 1985; Ma & Herasimchuk, 1975; Poplack, 1980; Terrell, 1978), but that deletion tends to occur more often when a vowel or pause follows (Cedergren, 1973). This shows that /s/-weakening is affected by the phonological environment (Bybee, 2000). Bybee (2000) discusses previous findings that Argentine and Cuban Spanish differed in how phonological environment affected /s/-weakening (Terrell, 1977, 1978, 1979). Specifically, Cuban Spanish was considered farther along in the change from [s]>[h]>[∅] because it showed much more aspiration and deletion in more innovative phonological positions, such as at the end of a word preceding a vowel. Other studies, such as File-Muriel (2007), have found that the manner of articulation of a following consonant affects /s/-weakening. In his study, /s/-weakening was more prevalent when followed by a fricative than a stop consonant in Columbian Spanish. Prosodic factors, too, can influence /s/-weakening. Brown (2005) showed a conditioning effect of stress in that syllable-initial /s/ was weakened more in unstressed syllables than stressed environments. File-Muriel (2007) did not find stress to be a significant factor in the realization of /s/, but other studies have found that when coda /s/ is in the stressed syllable it tends to be retained and is more likely to be weakened in atonic syllables (Alba, 1982; Brown, 2009; File-Muriel & Brown, 2011; Poplack, 1986). However, if followed by a stressed vowel, /s/ tends to be retained, which has been found in multiple Caribbean dialects (Cedergren, 1973; Terrell, 1978, 1981; Poplack, 1979).

The relationship between /s/-weakening and morphology has also been studied. Kiparsky (1972) proposed the distinctiveness condition of the functional hypothesis, which argued that semantically relevant information tends to be maintained in surface structure. Under the strong version of this hypothesis, it was predicted that [s] would not be deleted where it would cause ambiguity regarding the plurality of a word. Likewise, where multiple types of disambiguating information are present, a

phonological process such as deletion would be more likely to occur. Poplack (1980), Hochberg (1986), Hundley (1987) and Ranson (1991) all investigated this question. Their findings, though conflicting at times, generally showed that semantic information can reduce ambiguity upon the loss of /s/, such as when a quantifier is present that indicates plurality (Hochberg, 1986; Poplack, 1980). Hundley (1987) confirmed Poplack's (1980) finding that /s/ as a plural marker tends to be maintained if the inflectional marking in a noun phrase (NP) is the only disambiguating factor, but that deletion occurs the most when other morphological information within or outside the NP is available to compensate for the loss of [s]. He also found that deletion was likely to occur when other syntactic or semantic information was available to reduce ambiguity. Hochberg (1986) was the first to focus on the second person singular verb forms, which can be ambiguous in different tenses (e.g., present and imperfect, *yo comía* vs. *tú comía(s)*). She hypothesized that more overt subject pronouns would be used to compensate for the ambiguity and indeed found high use of subject pronouns in second person singular verb forms compared to other forms. However, Ranson (1991) found the opposite in that there was a very high rate of [s] deletion in Andalusian Spanish, while at the same time no increase in subject pronoun use with verb endings rendered ambiguous by [s] deletion. She, like Poplack (1980), argued that there is a tendency to retain semantically relevant information by means of linguistic or contextual information rather than surface morphological structure. What these studies on functional compensation show is that the retention or deletion of [s] is not constrained primarily by functional factors to retain plural information in the noun phrase, as was originally hypothesized. The deletion of [s] can instead be facilitated by the presence of other disambiguating factors such as contextual, semantic, and syntactic information. A recent study by Ruiz-Sánchez (2005) regarding the effect of functional factors on /s/-weakening in Caracas Spanish corroborated this finding that a strong version of the functional hypothesis is not warranted by the data. Despite findings that functional factors do not influence /s/-

weakening as much as originally assumed, there are other factors that have been found to play a strong role.

In addition to language-internal factors, the production of /s/ has been shown to be conditioned strongly by social and contextual factors. In general, there tends to be a social difference between aspiration and deletion, namely that aspiration is typically not as socially stigmatized as deletion (Calero, 1993; Carbonero, 1982; Carvalho, 2006; Cid-Hazard, 2003; Lynch, 2009; Ruiz-Sánchez, 2004; Valdivieso & Magaña, 1991), and may even be the norm in some dialects (Villena Ponsoda, 2008). Deletion, on the other hand, tends to be more common in the lower socioeconomic classes than the middle and upper classes (Alba, 2004; Carvalho, 2006; Cepeda, 1995; Lipski, 1984) and among those with a lower level of education (Alba, 1982; Lafford, 1986; Uber, 1989). The aspiration of /s/ can also be associated with lower socioeconomic class, but it tends to be more socially neutral than deletion, being common among middle and even upper class speakers in some dialects (e.g., Lafford, 1986; Villena Ponsoda, 2008). For example, Lafford (1986) found that aspiration of /s/ did not act as a clear social marker, lacking correlation with specific social classes. The deleted variant ([∅]), however, was associated more with the lower social classes. Nonetheless, socioeconomic class is not the only factor that plays a role. Lafford (1986) and others have also taken into account the formality of speech style, which can interact with socioeconomic class.

Speech style is an important factor in the realization of /s/ in that there tends to be more weakening and deletion in informal speech styles than formal styles (Carvalho, 2006; Cid-Hazard, 2003; Lafford, 1986). Lafford (1986), for example, found more full [s] in formal styles and higher social classes, while she found aspiration (i.e., [h]) to be neutral in conversational style but stigmatized in formal styles. The deletion of [s] was more common among the lower social classes in formal speech styles, but the gap between the upper and lower classes was lesser in informal speech styles. Younger speakers preferred [s] in formal styles, stigmatizing both [h] and [∅]. However, in informal styles, they deleted [s]

more than older speakers. Thus, the age of the speaker is a factor that can also affect /s/-weakening, but its effect has been inconsistent in studies to date, and has shown differences in how age affects aspiration versus deletion. Some studies have found that younger speakers favor /s/-weakening more than older speakers (Lipski, 1984; Lynch, 2009; Ruch, 2013; Ruch & Harrington, 2014). However, others have found younger and older generations to be similar (Lafford, 1986) or that older speakers weaken /s/ more than younger speakers (Alfaraz, 2000; Cedergren, 1973). This can depend on whether aspiration or deletion is the variant in question (Cedergren, 1973) or also on dialect, as Ruch and Harrington (2014) found age differences between Western and Eastern Andalusian Spanish speakers' treatment of /s/-aspiration. Finally, research has shown that the sex of the speaker is also a possible social predictor of /s/-aspiration and deletion given that some studies have found that /s/-weakening, and particularly deletion, is more common among men than women (Cedergren, 1973; Cepeda, 1995; Lafford, 1986) while in other studies it is not a strong factor (Alfaraz, 2000). Cedergren (1973) found a difference between aspiration and deletion with regard to speaker sex in that women favored the aspirated variant while men favored deletion, the more stigmatized variant.

In addition to sociolinguistic studies, some recent studies on Spanish /s/-aspiration and deletion have taken a usage-based approach (Bybee, 2002) to the study of how /s/ has changed over time and how this has been influenced by the frequency of lexical items. In this theoretical approach, every word experienced by the speaker is stored with detailed phonetic, contextual, and talker information, allowing for language change to take place gradually over time and spread from word to word as associations are made between them. Moreover, words and phrases used in higher frequency are more often phonetically reduced than low-frequency words (Bybee, 2002; Pierrehumbert, 2001). Multiple studies have found that /s/-weakening is significantly affected by lexical frequency (Brown, 2009; File-Muriel, 2007, 2009; File-Muriel & Brown, 2011). Specifically, coda /s/ in words with a higher lexical frequency tends to weaken more than in words with lower lexical frequency and these studies support

the hypothesis that /s/-weakening spreads gradually through the lexicon over time rather than all at once.

A few recent studies have begun to recognize the phonetically gradient nature of /s/-weakening and incorporate important acoustic measures into study design, removing the subjectivity of most previous studies that used impressionistic coding of /s/ tokens. This is important because studies such as Widdison (1995, 1997) have shown that /s/-aspiration involves a complex overlap of oral and glottal gestures in precise timing relationships. With this in mind, a study by Erker (2010) on /s/-weakening in Dominican Spanish used the acoustic measurements of duration and center of gravity as dependent variables. Erker found significant subsegmental variation between productions of /s/, rather than a tripartite distinction between [s], [h] and [∅]. In other words, using a segmental description puts tokens in the same group that are actually quite different acoustically. Similarly, File-Muriel & Brown (2011) included three acoustic measures as dependent variables: /s/-duration, centroid, and the percent of voicelessness. They found that an increased rate of speech led to more lenition and assimilatory processes than slower speech rates, and that this was the most significant predictor of /s/ productions. Other significant factors were word position, surrounding phonological context, stress, informant, and lexical frequency. The results of these studies that take into account gradient acoustic measures show the importance of accounting for variation in study design and the realization that second language learners are not exposed to discrete categories, but rather to a wide range of acoustic variation that they must learn to process and categorize correctly in the L2 phonological system. With this goal in mind, the following is a description of the acoustics and sociolinguistic variation surrounding /s/-aspiration in Andalusian Spanish in order to describe what will be in the input of L2 learners of Spanish who study abroad in Seville, Spain.

A Phonetic Characterization of Andalusian /s/-aspiration

In the coda position of a syllable in Spanish, there are multiple possible allophones of the phoneme /s/. The primary allophone, [s], is described as a voiceless alveolar fricative (Quilis, 1999), which is produced using high-frequency turbulent noise in the range of 6 to 10 kHz. In the spectrogram below, it can be seen that [s] exhibits a period of high frequency energy (i.e., the dark band on the spectrogram) that is accompanied by a lack of periodicity in the waveform, showing that it is voiceless.

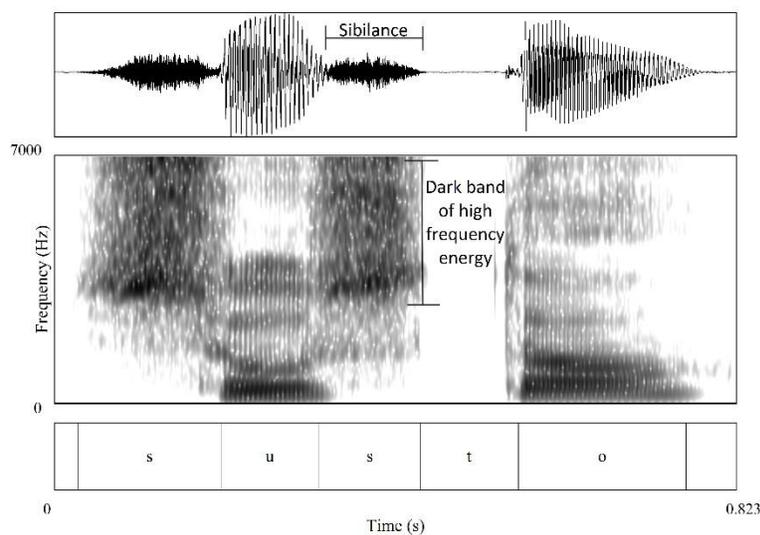


Figure 1: Example of a sibilant, *susto* [susto]

The aspirated allophone ([h]) has been described as the loss of all supralaryngeal features (i.e., debuccalization) while maintaining a spread glottis (Morris, 2000), producing voiceless frication without the high frequency noise that occurs in the production of the sibilant [s] (see “B” in Figure 2).

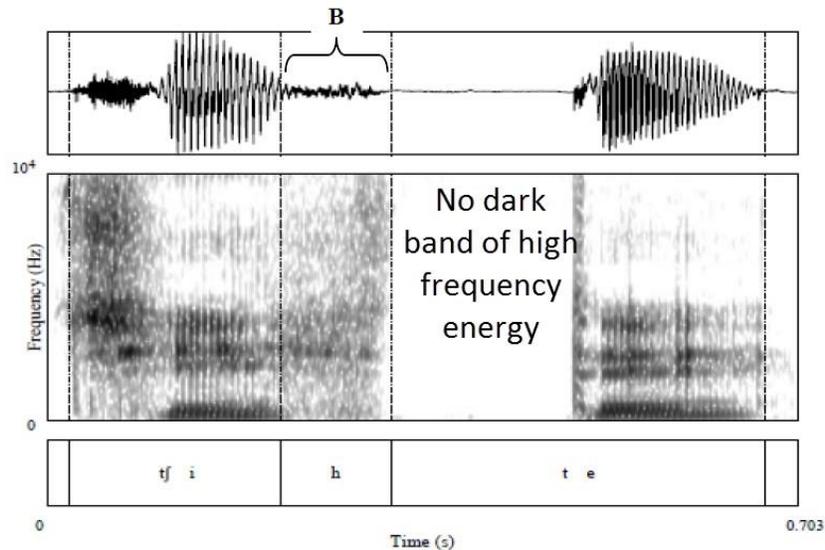


Figure 2: Example of aspiration from Schmidt (2011, p. 11) – *chiste* [tʃihte]

However, we have seen in the review of the literature that the typical IPA categorization [h] that has permeated the literature on /s/-weakening is too simplistic to represent the dialectal in the phonetic realizations of this allophone (Erker, 2010; File-Muriel & Brown, 2011). While in some dialects it is true that the aspirated variant of /s/ is realized as voiceless frication (i.e., [h]), in other dialects, such as Andalusian Spanish, /s/-aspiration can also be produced by means of different acoustic cues, particularly when preceding a voiceless stop.

The first acoustic cue of importance in /s/-aspiration in Andalusian Spanish is the realization of the aspiration preceding a voiceless stop. In a phonetic study on coda /s/-aspiration in four varieties of Andalusian Spanish (Seville, Cádiz, Almería & Granada), O'Neill (2010) showed that there were three ways of realizing the coda /s/: as a sibilant [s] (the least common), with glottal activity (i.e., breathy phonation, voiceless aspiration, or a combination of the two) and without glottal activity (i.e., as in words with no coda consonant, such as *pata* [pata] 'paw'). When there was glottal activity in place of the sibilant [s], voiceless strident aspiration such as what is found in places such as Argentina (Torreira, 2006) was not the most common realization in Andalusian Spanish. Instead, the glottal activity in

Andalusian /s/-aspiration was of a different type that tended to be voiced, less strident, shorter in duration and had a lower amplitude (i.e., breathy phonation) than the aspiration of most aspirating varieties of Spanish. Gerfen (2002) also attested to this realization. Furthermore, in Andalusian Spanish it is common for there to be no glottal activity at all after the first vowel and preceding the stop closure of the following voiceless stop (e.g., *pasta* ‘pasta/paste’ [pat^ha]). O’Neill (2010) termed this a ‘NOGLOT’ realization. This production was common in Seville, where it had an almost equal distribution with GLOT productions (i.e., productions with glottal activity, most frequently breathy phonation). In Cádiz, there was a preference for realizations with breathy phonation or a combination of breathy phonation and strident aspiration, but there was still evidence of the NOGLOT realization in 23 percent of cases.

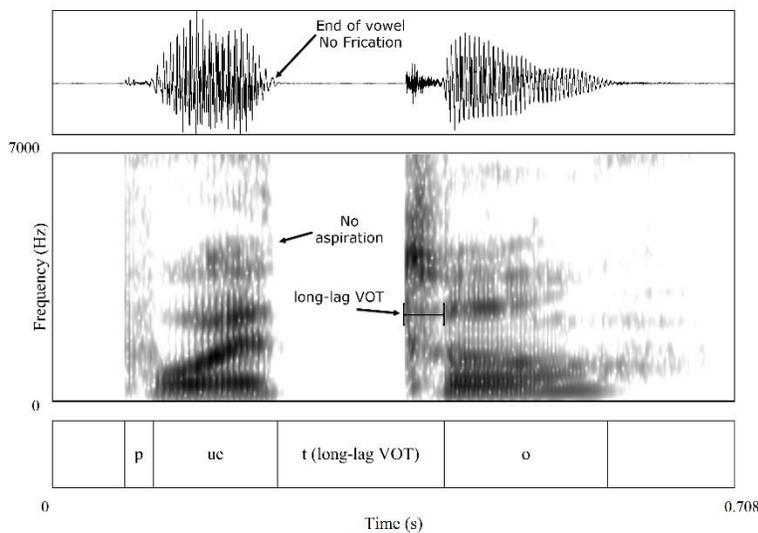


Figure 3: Example of NOGLOT with a post-aspirated /t/, *puesto* [pue^th^o]

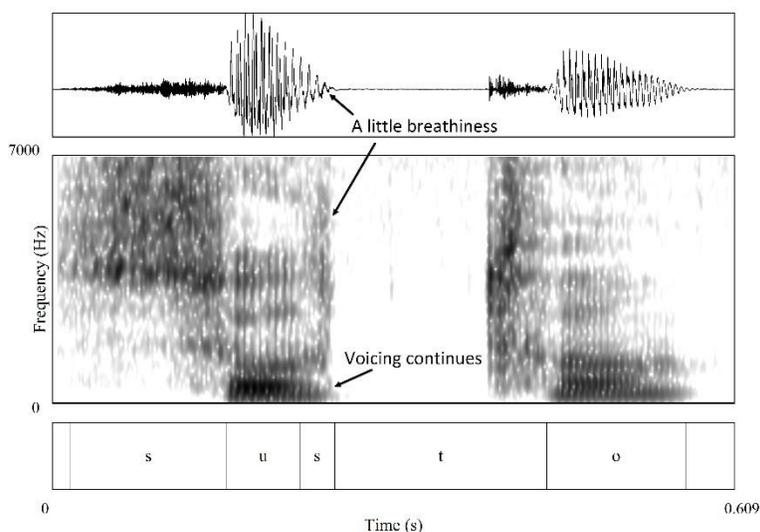


Figure 4: Example of breathy voicing with a post-aspirated /t/, *susto* [sutʰo]

Another very salient acoustic cue for Andalusian /s/-aspiration is the lengthening of the Voice Onset Time of the voiceless stop that follows /s/ in words such as *pista* ‘clue’ (i.e. long-lag VOT, henceforth referred to as ‘post-aspiration’). An example of a long-lag VOT in an aspirated word produced by an Andalusian talker be seen in Figures 3 and 4. Regular voiceless stops in Spanish exhibit a short-lag VOT of between approximately 13 milliseconds and 27 milliseconds, depending on which voiceless stop it is (Rosner et al., 2000). The shortest VOT tends to be for [p], while the longest tends to be for [k]. An example of a normal VOT for /t/ is shown in Figure 5. In dialects that produce aspiration ([h]) preceding a voiceless stop, such as Argentinian Spanish, the VOT of the voiceless stop following /s/ tends to be similar in duration to typical Spanish voiceless stop VOT durations (see Torreira, 2006 for Argentine and Puerto Rican Spanish VOT durations). However, Torreira (2006, 2007) showed that in Western Andalusian Spanish, long-lag VOT occurred for all three voiceless stops (/p t k/), and that this was the most prominent acoustic cue for Andalusian /s/-aspiration. Importantly, the long-lag VOT in the /CVS.CV/ context differs significantly from the VOT of voiceless stops in /CV.CV/ sequences such as *pata*

'paw' (i.e., [pata]), which has short-lag VOT (O'Neill, 2010: 25; Torreira, 2006). In addition to the findings of Torreira (2006) and O'Neill (2010), Ruch's (2013) study on phonetic and lexical factors related to /s/-aspiration in Andalusian Spanish found two different types of productions of the voiceless stops following a coda /s/. She found a long-lag VOT production like what was described by O'Neill and Torreira (e.g., [t^h]), also an affricate stop realization in the case of /st/ clusters (i.e., [t^s]). This type of realization was also found by Moya Corral (2007).

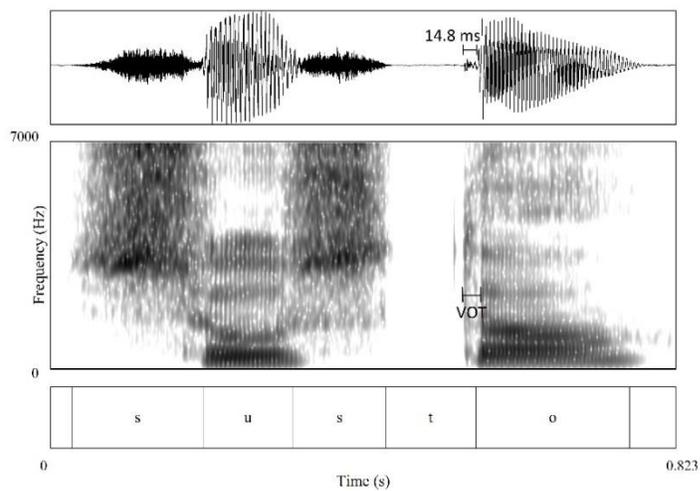


Figure 5: Example of a regular /t/ (not post-aspirated), *susto* [susto] (14.8 ms VOT)

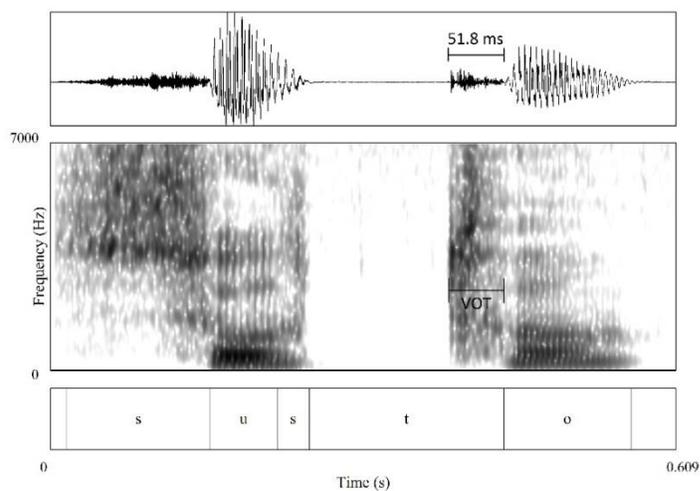


Figure 6: Example of a post-aspirated /t/, *susto* [sut^ho] (51.8 ms VOT)

In Andalusian Spanish there are additional phonetic cues for /s/-aspiration that can occur in /VS.CV/ sequences, such as stop-closure lengthening (Gerfen, 2002; Torreira, 2006) and a longer vowel preceding /s/ than the first vowel in /CV.CV/ sequences (Gerfen, 2002; O'Neill, 2010; Torreira, 2006). However, as Gerfen (2002) noted, the finding that vowels are longer in /CVS.CV/ sequences than /CV.CV/ sequences is really only significant if aspiration in coda position is considered part of the vowel (i.e., /CV[h].CV/). The duration of the vowel in /CVS.CV/ sequences without including the duration of the aspiration tends to be shorter than the vowel in /CV.CV/ sequences, as would be expected given the universal fact that vowels are generally shorter in closed syllables. Eastern Andalusian Spanish exhibits further variation related to vowels in that /s/-weakening tends to cause vowel harmony as a compensatory mechanism (Gerfen, 2002). However, this has not been shown for Western Andalusian Spanish.

Regarding the length of the stop closure preceding the release burst of the voiceless stop following /s/ in /VS.CV/ sequences, Torreira (2006) and Ruch and Harrington (2014) found the duration of the stop closure to be significantly longer in post-aspirated sequences (e.g., *pasta* 'pasta' [pat^ha]) than in /V.CV/ sequences (e.g., *pata* 'paw' [pata]). In comparison, there was no significant difference in closure duration between /VS.CV/ and /V.CV/ contexts for the Puerto Rican or Buenos Aires speakers. Torreira's (2006) data were collected from three groups of native speakers of Spanish from Western Andalusia, Buenos Aires and Puerto Rico in a laboratory setting. The speakers read words embedded in the carrier sentence *Digo _____ para mí* 'I say __ for myself.' Ruch and Harrington's (2014) data come from 24 Eastern Andalusians and 24 Western Andalusians who produced four isolated target words with word-medial /st/ sequences and two words with medial /t/ in intervocalic position. Each of the words was repeated three times. There were 130 items with intervocalic stops or stop clusters that served as filler items. O'Neill (2010), in his study of /s/-aspiration in Eastern and Western Andalusian Spanish that included laboratory speech that was recorded by eight subjects from Seville (N=2), Cádiz (N=2), Granada

(N=2) and Almería (N=2) that read words embedded in the carrier phrase *dame un _____ pa mí* ‘give me a _____ for me’, only found a significant difference between the closure durations for /CVS.CV/ and /CV.CV/ sequences for Granada and Almería (i.e., Eastern Andalusia) but not Seville or Cádiz (i.e., Western Andalusia). Thus, according to these three lab studies, while the duration of the stop closure appears to be a cue for Andalusian /s/-aspiration, it is not as consistent as VOT and the duration of aspiration. However, future research should incorporate spontaneous speech data for comparison to the laboratory data in order to determine if the laboratory data is truly representative of spontaneous speech.

As a summary of some of the data from the studies to date on Andalusian /s/-aspiration, Table 2 displays the mean durations of VOT and stop closures in Seville and Cádiz (i.e., Western Andalusian varieties) for /VS.CV/ and /V.CV/ sequences as reported by O’Neill (2010) and Torreira (2007)². Tables 4 and 5 present the comparisons of VOT and stop closure duration among old and young speakers of Eastern Andalusian Spanish (EAS) and Western Andalusian Spanish (WAS) as reported by Ruch and Harrington (2014).

² I generated the mean VOTs from Torreira (2007) by calculating the overall mean of the average VOT durations of the three speakers who were listed as being from Seville. The tokens came from recordings of interviews from an Andalusian public television channel and represented spontaneous data. Mean VOT and closure measurements for laboratory results from Torreira (2006, 2007) were not explicitly given. One must read the graphs for the results, which do not provide information regarding exact means. However the graphs in the 2006 paper show that the mean VOTs of [p], [t] and [k] in WAS /VS.CV/ words were approximately 30ms, 35ms, and 52ms respectively, with a large amount of variation ranging from approximately 15ms on the low end to approximately 80ms on the high end. In contrast, VOTs for Andalusian stops in /CV.CV/ contexts ranged from approximately 5ms on the low end and 22-25ms (for /p/ and /t/) and 44ms (for /k/) on the high end. The 2007 paper shows Andalusian /VS.CV/ VOT ranges between approximately 15-55ms for /p/, 18-75ms for /t/ and 35-78ms for /k/ with approximate means of 30ms, 35ms and 55ms respectively. In contrast, the standard Spanish speaker group (i.e., Northern Peninsular Spanish) produced VOTs ranging between approximately 5-22ms for /p/, 5-30ms for /t/, and 15-40ms for /k/ with approximate means of 10ms, 12ms, and 28ms respectively.

Table 2. Mean durations of VOT and stop closures in Seville and Cádiz (Western Andalusia) for /VS.CV/ and /CV.CV/ sequences (O'Neill, 2010 and Torreira, 2007)

Variety of /s/-aspiration	VOT (VS.CV – e.g., [pat ^h a])	VOT (V.CV – e.g., [pata])	Stop closure duration (VS.CV)	Stop closure (V.CV– e.g., [pata])	Speech type
Seville (O'Neill)	51.2 ms	15.81 ms	60.69 ms	58.17 ms	Lab speech (N=2)
Seville (Torreira)	40.78 ms	16.17 ms	n/a	n/a	Lab speech (N=9)
Cádiz (O'Neill)	62.8 ms	21.86 ms	45.35 ms	63.21 ms	Lab speech (N=2)

Table 3. Statistics for the duration of coda glottal activity (i.e. GLOT) in both cities of Western Andalusia from O'Neill (2010)

Variety of /s/-aspiration	Mean duration of glottal activity	SD	Speech type and number of speakers
Seville	14.97 ms	8.96 ms	Lab speech (N=2)
Cádiz	22.45 ms	11.22 ms	Lab speech (N=2)

Table 4. Statistics for mean VOT in /VS.tv/ contexts compared to /V.tv/ contexts, comparing older to younger speakers in East and West Andalusia (Ruch & Harrington, 2014).

Variety of /s/-aspiration and syllable structure	Younger Mean VOT	SD	# of speakers	Older Mean VOT	SD	# of speakers
EAS /VS.tv/	35.8 ms	14.0 ms	12	20.4 ms	8.1 ms	12
EAS /V.tv/	18.7 ms	6.4 ms	12	16.3 ms	7.5 ms	12
WAS /VS.tv/	55.2 ms	17.5 ms	12	26.8 ms	10.3 ms	12
WAS /V.tv/	18.7 ms	7.4 ms	12	19.2 ms	11.5 ms	12

Table 5. Statistics for mean stop closure duration in /VS.tv/ contexts compared to /V.tv/ contexts, comparing older to younger speakers in East and West Andalusia (Ruch & Harrington, 2014).

Variety of /s/-aspiration and syllable structure	Younger Mean SC	SD	# of speakers	Older Mean SC	SD	# of speakers
EAS /VS.tv/	113.4 ms	26.2 ms	12	115.2 ms	23.8 ms	12
EAS /V.tv/	80.4 ms	17.1 ms	12	81.0 ms	18.5 ms	12
WAS /VS.tv/	88.7ms	20.3 ms	12	110.0 ms	22.3 ms	12
WAS /V.tv/	68.9 ms	16.9 ms	12	81.1 ms	17.7 ms	12

In a recent study, Torreira (2012) analyzed Andalusian /s/-aspiration from the perspective of Articulatory Phonology (Browman & Goldstein, 1992) to determine the cause of long-lag VOT in productions of aspirated /st/, /sp/ and /sk/ clusters in Andalusian Spanish. Based on his findings he

proposed that long-lag VOT in /VS.CV/ contexts in Andalusian Spanish resulted from a change in the gestural timing relationship between glottal spreading for the coda /s/ aspiration gesture and the oral constriction of the following voiceless stop. His findings suggest that a change from sequential timing (i.e., glottal spreading followed by the oral stop closure) to simultaneous timing of the two gestures resulted in the long-lag VOT.

Parrell (2012) hypothesized that speech rate might play a role in the gestural alignment relationship described by Torreira (2012). He recorded /CVS.CV/ sequences in Andalusian Spanish at increasing rates of speech using click timing intervals. He hypothesized that at slow rates of speech Andalusian speakers would produce more glottal activity prior to the voiceless stop and normal VOTs, but that faster speech rates would lead to little to no glottal activity prior to the stop concomitant with a long-lag VOT. Intermediate speech rates were hypothesized to produce intermediate values for both acoustic cues. His results supported the hypothesis, finding that long-lag VOT resulted from gestural realignment that was affected by rate of speech. The productions went from short to long-lag VOT as speech rate increased and there was an inverse relationship between VOT duration and the duration of glottal activity preceding the stop closure. As VOT increased, the glottal activity decreased. Still, there were some cases in which increasing rate of speech did not change VOT from short to long, and some long-lag VOTs were even produced at slow rates of speech, showing that the phase relationship between the two gestures is not the only factor. Parrell (2012) argued that in the grammar, these consonant clusters are stored as /s/+stop and not the other way around, which means that gestural reorganization is being resisted by the grammar and thus may not always occur. He also suggests that language-external factors are at play that may influence speakers, such as sociolinguistic differences between Andalusian Spanish and Northern Peninsular Spanish, which does not allow for this gestural reorganization. Given that the research occurred in a laboratory setting with a non-native researcher, he hypothesized that there may have been an influence of these factors.

Ruch and Harrington (2014) agree that the timing relationship between the glottal and oral gestures cannot completely account for long-lag VOT productions. In fact, they found some opposite trends compared to Parrell (2012) and Torreira (2012). Specifically, they found in the productions of 24 speakers from Seville (WAS) and 24 from Granada (EAS) that there was no significant relationship between the duration of aspiration (i.e., GLOT) and post-aspiration (i.e., VOT) for the stop /t/. Additionally, the two varieties, WAS and EAS, differed in the length of post-aspiration but not aspiration, indicating that aspiration and post-aspiration are independent. Thus, they argue that these findings do not support Parrell's (2012) result that a shorter aspiration led to a longer post-aspiration and vice versa. Ruch and Harrington (2014) also found that younger speakers produced shorter stop closure durations than older speakers in /st/ words but not /t/ words, and that there was a relationship between the length of the stop closure and the length of post-aspiration for younger speakers and between length of closure and aspiration (i.e., glottal activity preceding the stop) for older speakers, demonstrating that the lengthening of the stop closure (i.e., gemination) came about before long-lag VOT in these varieties. They suggest that future physiological and perceptual studies might find that VOT lengthening came about due to a long geminate closure that caused air pressure to rise and lead to a longer release burst. They also hypothesize that the trading relationship between closure duration and VOT for younger speakers suggests that if long-lag VOT continues to get longer and the closure duration shorter over time, VOT "may eventually become the primary cue for distinguishing /st/ from intervocalic /t/ in Andalusian Spanish" (p. 24).

Ruch and Harrington (2014) also conducted the first study testing the perception of Andalusian /s/-aspiration by native speakers of Spanish. They synthesized two continua between /pasta/ and /pata/ that included nine durations of /s/-aspiration, as well as one continuum with a long-lag VOT (post-aspirated) and another with a normal VOT of 12 milliseconds. They tested Argentinian listeners, who do not post-aspirate but do aspirate (e.g., [pahta]). While in Andalusia there is evidence of a sound change

from aspiration before the stop to post-aspiration (Parrell, 2012; Ruch & Harrington, 2014; Torreira, 2012), this is not true for Argentinian Spanish (Torreira, 2006). The purpose of the perceptual test was to determine if the change from aspiration to post-aspiration was based on perception. They took a stimulus from a WAS speaker that had an aspiration duration of 57 ms, a VOT (i.e., post-aspiration) of 45 ms and a closure duration of 139 ms and modified it for the continua. After modifications, there were two continua. Both had a closure duration of 100 ms. The short continuum had a VOT of 12 ms and the long continuum had a VOT of 29 ms. Both continua had aspiration durations that ranged from 0 to 34 ms in nine increments. The subjects were 64 NSs of Argentinian Spanish (39 women and 25 men) between the ages of 13 and 64. They were asked to listen to the words and determine if it was 'pasta' or 'pata' by clicking one of two choices. The results showed that for stimuli with a short aspiration, there were more 'pasta' responses when a long VOT was present, showing that VOT played a significant role. The longer aspiration (i.e., [pah.ta]) became, the less of a role the long VOT played in their responses. The results show that even for non-post-aspirating speakers, long-lag VOT was a cue for /st/, showing that "the emerging post-aspiration is parsed with pre-aspiration³ and is perceptually associated with the underlying phonological /st/" (p. 23).

In summary, the studies to date show a unique pattern for /s/-aspiration in Andalusian Spanish that seems to be a change in progress that has resulted from phonetic factors that could be dialect-specific gestural reorganization between the glottal and oral gestures or pre- and post-aspiration (Parrell, 2012; Torreira, 2012) or a relationship between a shortening stop closure and longer VOT (Ruch

³ What Ruch & Harrington (2014) refer to as 'pre-aspiration' is referred to in this study simply as 'aspiration,' referring to glottal activity preceding the stop closure of a following voiceless stop. In fact, aspiration has been referred to as 'pre-aspiration' in most studies to date. However, this is an erroneous categorization of what is actually occurring at the phonetic level. Aspiration that precedes the voiceless stop is not 'pre-' aspiration in the same sense that long-lag VOT is 'post-' aspiration. Post-aspiration is part of the gestural score of the stop itself, while aspiration preceding the stop is not part of the stop, but is a separate segment. This is demonstrated by the fact that /s/-aspiration can occur in other environments than before a voiceless stop, separating it from the gestural score of the stop itself. In this way aspiration is not the same as a pre-aspirated stop, which do occur in some languages, though infrequently (Silverman, 2003). Therefore, aspiration in Spanish should not be labeled 'pre-aspiration.'

& Harrington, 2014). Whatever the case, /s/-aspiration in WAS is variable and it has been shown that it can be perceived even by NSs from a dialect that does not exhibit long-lag VOT in /VS.CV/ contexts. What has yet to be determined in the literature is what effect this has on the perception of L2 learners, specifically how learners categorize these acoustic cues and what effect they have on lexical access. Learners who are exposed to the acoustic cues of Andalusian /s/-aspiration would receive input that is particular to this region of the Spanish-speaking world, making students who are exposed to Western Andalusian Spanish during study abroad a particularly interesting case when studying the perception of /s/-aspiration in Spanish. However, given that this phenomenon is sociolinguistically variable, it is important to explore its prevalence in the region as well as its social stratification based on research to date.

Variation of /s/ in Andalusian Spanish According to Social and Contextual Factors

Though very few studies to date have investigated the sociolinguistics of /s/-aspiration in Andalusian Spanish, there are a few that can provide some context as to how /s/-aspiration is viewed and used in Andalusian Spanish. Villena Ponsoda (2008), in describing /s/-aspiration in Andalusian Spanish, notes that the retention of [s] is generally rare in this variety and that aspiration frequently occurs even in the speech of the most educated speakers and in the most formal speech styles, leading Ruch (2013) to call aspiration in this dialect a “local standard” (p. 171). The most favoring environment for aspiration in Andalusian Spanish is word-medial before a consonant and, importantly, Seville is considered the regional center of Western Andalusia and is the model for pronunciation in this region (Villena Ponsoda, 2008).

To analyze the effect of language-external factors on /s/-aspiration in Andalusia, Moya Corral (2007) studied /s/-aspiration in Seville from a sociolinguistic perspective. The participants were NSs from the city of Seville or the nearby town of Antequera and they completed a sentence reading task, a

picture description task, and a word list reading task. He found that the aspirated variant in which a voiceless [t] following /s/ (i.e., [t^h]) had a lengthened VOT was much more common in informal speech than any other variants of /s/. He also found that younger speakers and speakers with a secondary level of education preferred this variant more than others. He did not find any significant differences between men and women. He concluded that the aspirated [t^h/p^h/k^h] variant is a prestige variant in the city of Seville, but not in Antequera.

Ruch (2013) studied the effects of phonetic and lexical factors on /s/-aspiration in /st/ clusters in Andalusian Spanish, while also accounting for age differences between the participants and differences between Eastern and Western Andalusia. Her results showed that among the 48 participants (24 from Seville and 24 from Granada), the younger speakers produced shorter duration of glottal activity preceding the stop closure and longer VOTs than the older speakers, indicating a change in progress from aspiration to post-aspiration, especially in Western Andalusian Spanish (i.e., Seville). She also found other statistically significant factors such as phonological context and lexical frequency. Ruch (2013) showed that /s/-aspiration in this variety is likely “the result of a gradual process in which post-aspiration gradually increases while pre-aspiration decreases” (p. 179). In this variety she shows overall that /s/-aspiration is, “of a regular more than of a sporadic nature” (p. 179), in agreement with Villena Ponsoda (2008), indicating that L2 learners exposed to this variety should be easily exposed to /s/-aspiration.

Now that the target variant, /s/-aspiration in Andalusian Spanish, has been thoroughly described both in terms of its acoustic properties and sociolinguistic patterns, it is necessary to discuss theories of L2 phonology and hypotheses related to how learners might process /s/-aspiration when exposed to it.

Second Language Phonology

In order to be able to hypothesize how L2 learners might perceive Western Andalusian /s/-aspiration and to provide the foundation for the methodological design of the current study, it is necessary to explore the field of second language phonology and theories of L2 speech perception as well as the merits of the two perception task paradigms that will be employed in the methodology of the current study. Second language phonology (henceforth 'L2 phonology') combines the fields of second language acquisition, phonology and psycholinguistics. Phonology is the branch of linguistic inquiry that is concerned with the organization of sound systems of languages, meaning the relationships among the sounds of a language and the role that each sound plays in the structure of a language. Researchers in the field of L2 phonology study aspects of the phonological system such as phonemic contrast, phonetics, allophonic variation and suprasegmental phenomena (i.e., intonation, stress, tone and pitch accent), but with a focus on how they are related to the developing phonological system of second language learners. Many researchers have focused on how L2 learners *produce* speech sounds in the second language, including research on the articulatory properties of L2 speech production, for example. Others researchers have focused their efforts on the issue of foreign accent and what factors best explain a low or high degree of foreign accent in L2 learners, such as age of acquisition and motivation (Flege, Piske & McKay, 2001). And finally, other research focuses on speech perception, which is the study of how an acoustic signal is mapped onto phonemic categories, and word recognition, which is the manner in which entire words are processed from the acoustic signal. Speech perception and word recognition fall under the domain of 'psycholinguistics.' While psycholinguistics encompasses many things, one goal is to explain how speech sounds and words are processed from an acoustic signal. The aims of the current study include an investigation of both speech perception (i.e., categorization) and word recognition (i.e., lexical access).

Before discussing L2 speech perception it is important to begin with the foundations of human speech perception early in life. Research on how infants discriminate sound contrasts in different languages has shown that, even before their first birthday, infants go from possessing a universal ability to discriminate sound contrasts in any language to a language-specific discrimination ability, being able to only discriminate contrasts in the L1 to which they are exposed (Kuhl, 2000; Werker & Tees, 1984; Polka & Werker, 1994). The implication for L2 speech learning is obvious, that the learning of L2 speech sounds must be influenced in some way by having language-specific perception based on the L1 phonological inventory so early in life. Proposals regarding the nature of the relationship between the L1 and L2 phonological systems, particularly the influence of the L1 phonology on the L2 phonology have been a central aspect of theories of L2 phonology.

In the early days of L2 phonological research it was hypothesized that L2 learners processed L2 phones via the L1 phonological system (Polivanov, 1931; Trubetzkoy, 1939/1969). In other words, they hypothesized that L2 learners would perceive and produce L2 phonemes as if they were L1 phonemes, which implies that the L1 phonological system affects, both positively and negatively, the developing L2 phonological system depending on the relationship between L1 and L2 phones. This led to hypotheses such as the Contrastive Analysis Hypothesis (Lado, 1957), which proposed that L2 phones that have a counterpart in the L1 should be easy to acquire and those that have no L1 counterpart should be difficult to acquire. While in some cases this hypothesis has been shown to have some explanatory power, such as for Japanese speakers' poor discrimination of English /r/ and /l/, due to Japanese not having an equivalent to the English /l/ (Goto, 1971; Best & Strange, 1992), this model has been shown to be too simplistic to account for L2 speech perception generally because not all segmental contrasts are equally difficult and not all L1-L2 combinations cause equal difficulty (Best, Hallé, Bohn & Faber, 2003; Best & Strange, 1992; Best & Tyler, 2007; Hallé, Best & Levitt, 1999). Nevertheless, the basic assumption that the relationship between the L1 and L2 phonological systems is important, and that the L1

influences L2 underlies other hypotheses that have been put forth about the influence of phonological forms in the L1 and L2 during acquisition, phonological or otherwise, such as the Native Language Magnet Model (Kuhl, 1992, 1993, 1994), the Speech Learning Model (Flege, 1995), the Perceptual Assimilation Model (Best, 1994, 1995; Best & Tyler, 2007), the Phonological Interference Model (Brown, 2000), the Ontogeny-Phylogeny Model (Major, 2001), and the Second Language Linguistic Perception Model (Escudero, 2005). However, only two of these models will be discussed and applied to the current study because most of the models either have a focus on production rather than perception, or have their base in theoretical assumptions that do not guide the design and implementation of the current study (e.g., Escudero, 2005 – Optimality Theory).

Models of L2 Speech Perception

The two primary models of L2 speech perception in the literature have been the Speech Learning Model (Flege, 1995) and the Perceptual Assimilation Model (Best, 1994, 1995; Best & Tyler, 2007). Flege (1995) proposed the Speech Learning Model (SLM) with the purpose of accounting for how individuals learn to or do not learn to produce and perceive phonetic segments in a second language, and how their abilities change across the lifespan, allowing for the creation of new phonetic categories during L2 learning. The SLM assumes that L1 acquisition processes and mechanisms remain intact and can be used by the individual for the duration of life, though it also predicts that while late learning is not impossible, early learning is easier than late learning. Specifically, Flege argued based on evidence from a variety of studies that there is no sharp cutoff for language learning after a particular age, as Lenneberg (1969) proposed in the Critical Period Hypothesis, but that there is more of a gradual decline over time. The SLM also counters earlier held positions that L2 sounds are perceived as L1 phonemes regardless of phonetic differences (Polivanov, 1931; Trubetzkoy, 1939/1969), which would imply that phonetic differences are discarded during processing. Instead, he argues that with increasing target

language experience adults can detect cross-language phonetic differences and that these differences are stored in long-term memory representations. Additionally, in the SLM, L1 and L2 phonetic systems share a common phonological space and can influence one another, rather than being completely separate systems.

The SLM's primary hypothesis predicts that for sounds in the L2 that are similar to an L1 counterpart on a position-sensitive allophonic level it will be difficult for the listener to create a new category for that L2 sound. What the SLM means by a 'position-sensitive allophonic level' is that when judging the similarity between L1 and L2 sounds, the phonotactic constraints of the two languages are not ignored. Thus, two sounds will only be similar if they occur in the same allophonic context. For example, if a L2 sound occurs in syllable-final position, but the closest L1 sound, phonetically-speaking, occurs in word-initial position and not word-final position, there would be no 'equivalence classification,' meaning that they would not be considered similar. If a L1 and L2 sound are considered similar at the position-sensitive allophonic level, the L1 and L2 phonetic categories will merge, making discrimination difficult. On the other hand, if sufficient phonetic differences are detected with increased language experience, a new category will be created and the creation of a new category may cause the L2 sound to become less like the L1 and/or L2 sound in order to maintain phonetic contrast. In other words, the greater the phonetic distance between an L2 and an L1 sound, at the position-sensitive allophonic level, the more likely it is that a listener will detect the differences and create a new phonetic category. In the SLM, age plays an important role in that Flege hypothesized that as an individual ages he or she will require more phonetic distance between the L1 and L2 sound in order to form a new category. The SLM claims that problems with speech production in the L2 can be attributed to inaccurate perception. All of the postulates and hypotheses of the SLM are presented in Table 6.

Table 6. Postulates (P) and Hypotheses (H) of Flege's SLM (1995, p. 239)

- P1** The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span, and can be applied to L2 learning.
- P2** Language-specific aspects of speech sounds are specified in long-term memory representations called phonetic categories.
- P3** Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category.
- P4** Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space.

- H1** Sounds in the L1 and L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level.
- H2** A new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds.
- H3** The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that the phonetic differences between the sounds will be discerned.
- H4** The likelihood of phonetic differences between L1 and L2 sounds, and between L2 sounds that are non-contrastive in the L1, being discerned decreases as [Age of Learning] increases.
- H5** Category formation for an L2 sound may be blocked by the mechanism of equivalence classification. When this happens, a single phonetic category will be used to process perceptually linked L1 and L2 sounds (diaphones). Eventually, the two diaphones will resemble one another in production.
- H6** The phonetic category established for L2 sounds by a bilingual may differ from a monolingual's if: 1) the bilingual's category is "deflected" away from an L1 category to maintain phonetic contrast between categories in a common L1-L2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's.
- H7** The production of a sound eventually corresponds to the properties represented in its phonetic category representation.

The second model of nonnative speech perception is the Perceptual Assimilation Model (PAM: Best, 1994, 1995). This model was not originally created to explain L2 speech perception, but rather how naïve listeners perceive nonnative phones and phonemic contrasts. The PAM, based on Articulatory Phonology (Browman & Goldstein, 1992), proposed that segments are constellations, or clouds, of gestures and that individuals perceive non-native segments "according to their similarities to, and discrepancies from, the native segmental constellations that are in closest proximity to them in native phonological space" (Best, 1995: 193). In this way, non-native phones are assimilated to native language phonological categories and Best (1995) outlines predictions about how L2 phones would be

categorized, which has consequences for discrimination and acquisition, assuming that good discrimination should enable acquisition to occur (Table 7).

Table 7. PAM Categories of Assimilation (Best, 1995)

Two-Category Assimilation	Two non-native segments are assimilated to two separate native categories. Discrimination should be good.
Category-Goodness Difference	Non-native phones assimilate to one native category, and one phone is a better example of that native category than the other. Discrimination should be moderate to good.
Single-Category Assimilation	Two non-native phones assimilate to the same native category but do not differ in their goodness-of-fit to that category. Discrimination will be poor.
Uncategorizable	Non-native phones are not assimilated to any native category but are still within the phonetic space. Discrimination depends on how close the two phones are to one another and to native categories.
Non-Assimilable	Both non-native phones are not within the L1 phonetic space. Discrimination should be good to very good.
Uncategorized vs. Categorized	One phone is assimilated to a native category and the other is not, but still falls within the phonetic space. Discrimination should be good or very good.

Since the PAM was not created to account for L2 speech perception, but rather the perceptual patterns of naïve listeners, Best and Tyler (2007) slightly modified it for second language learners (PAM-L2), but with the same basic underlying framework as the original PAM. PAM-L2 assumes that the L2 learner begins like the naïve learner in the PAM, but progresses beyond that level of perception over time. The reason for this is that the L2 learner and the naïve listener will likely differ in multiple ways. For example, an L2 learner will likely attempt to produce target contrasts, learn lexical items in the target language and “re-phonologize perception of the target contrasts” (Best & Tyler, 2007, p. 32) while the naïve listener does not have motivation to do these things. An important difference between PAM-L2 and the SLM is that PAM-L2 differs from the SLM’s concept of equivalence classification in that the

PAM-L2 involves equivalence not only at the phonetic level but also at the phonological level. Because of this, two sounds may be phonetically dissimilar but be perceived as phonotactically and phonologically similar. The opposite may also be true, that two sounds may be phonetically similar but phonotactically and phonologically dissimilar. Best and Tyler (2007) give an example of the French /r/ ([ʁ]) and English /r/ ([ɹ]), which are similar phonologically, but different phonetically. According to the PAM-L2 they should be easily discriminated at the phonetic level and the L2 learner should be able to form a second phonetic category for the French rhotic as part of the common phonological category /r/, which is a shared phonological space between the two languages.

Application of L2 Perception Models and Theories to the Current Study

In view of the L2 perception models and theories presented above, /s/-aspiration in Western Andalusian Spanish presents an interesting case for L2 learners. As described above, in the production of Western Andalusian /s/-aspiration, the duration and quality of aspiration preceding a word-internal voiceless stop tends to be short, often exhibiting qualities of breathy phonation or no glottal activity at all. Additionally, the VOT of the voiceless stop following a weakened or elided coda /s/ tends to be long (O'Neill, 2010; Torreira, 2006, 2007, 2012). Given these phonetic facts, the following is a comparison of the predictions that the PAM and SLM make regarding how the phonetic cues of Western Andalusian /s/-aspiration might be perceived by English-speaking learners.

The first cue of interest is aspiration that precedes a voiceless stop, as in [CV[h].CV] sequences. A few hypotheses can be made about how this cue could be processed by learners of Spanish. First, if aspiration takes on a strident voiceless character, the PAM would predict assimilation to the English /h/ category given the similarity in the gestural scores of the two phones despite the fact that Spanish [h] and English /h/ do not follow the same phonotactic rules. Spanish [h] occurs, in most dialects that aspirate /s/, in the coda position of a syllable, while English /h/ occurs only in onset position (Giegerich,

1992). Likewise, a maintained [s] would be predicted to assimilate to the English /s/ category, creating a two category assimilation pattern for these two phones, leading to good discrimination between [s] and [h] on the phonetic level. However, the SLM does not make quite the same prediction for the Spanish [h] because it is not equivalent to English /h/ at the position-sensitive allophonic level. Therefore, the SLM predicts that [h] would not be considered equivalent to any English sound in the same word position and a new category would eventually be created with increased experience. Learners who are exposed to this type of aspiration should be able to improve their perception over time through use of the new category, as was shown by Schmidt (2011), who found that increased language experience and study abroad experience in a /s/-weakening dialect region led to more accurate mapping of [h] to orthographic <s>. Initially, though, it could be that the phonotactic constraints of English prevent the perception of [h] because of the illegal position according to English phonotactics. Dupoux et al. (1999), for example, showed that phonotactic knowledge can cause listeners to repair illegal structures to perceive only legal ones, in this case, a /CV.CV/ sequence rather than /CVs.CV/.

Still, /s/-aspiration in Andalusian codas is not as simple as strident voiceless aspiration preceding a voiceless stop. Most of the time aspiration in this dialect has been shown to be breathy phonation that is shorter than strident aspiration, and is often produced without any aspiration at all (O'Neill, 2010). This means that the predictions described above do not always apply to Andalusian /s/-aspiration. English-speaking learners of Spanish may perceive short breathy phonation as an extension of the voicing of the vowel and assimilate it to the native vowel category (PAM). In this case it would be hypothesized that learners would perceive these /CVs.CV/ sequences as /CV.CV/, if not taking into account the additional VOT cue in Andalusian Spanish. However, findings that have shown that word-medial vowels preceding voiceless consonants in American English tend to be shorter than vowels preceding voiced consonants (Eilers, Bull, Oller & Lewis, 1984) suggest that the SLM would predict that a seemingly longer vowel in aspirated sequences in Spanish would not assimilate to a long vowel

allophone in English if a phonetic difference is detected because vowel lengthening in English does not tend to occur before voiceless stops. Thus, the SLM would predict that this additional breathy voicing at the end of the vowel might eventually lead to the creation of a new phonetic category. If the breathy phonation is not initially perceived as being phonetically different than an English vowel, it will not be processed as a new phonetic category and /CV.CV/ structure would be perceived rather than /CVs.CV/. These predictions, though, ignores the fact that Andalusian /s/-aspiration is a multi-cue phenomenon, with long-lag VOT being the primary acoustic cue. Therefore making predictions solely based on one acoustic dimension is not truly representative of what occurs at the phonetic level. This highlights the importance of phonetic cue weighting, which should be a future topic of research on the L2 acquisition of /s/-aspiration in Spanish to determine how L2 learners weight the multiple acoustic cues for /s/-aspiration and whether cue weighting changes over time. Such studies would have to manipulate stimuli across multiple acoustic dimensions to determine the effects of each acoustic cue independently on the perception of L2 learners.

In order to predict how speakers of American English will perceive the long-lag VOT in Andalusian coda /s/-aspirated sequences, it is necessary to compare VOT in Spanish and English. In American English, the VOT of voiceless stops is significantly longer than that of Spanish voiceless stops in word-initial position (Hualde, 2005). Lisker and Abramson (1967) showed that English voiceless stops in word-initial position, particularly when beginning a stressed syllable, exhibited long-lag VOT. In word-internal onset position, though, the VOT of voiceless stops in English does not tend to be as long as in word-initial position (Lisker & Abramson, 1967). Based on this comparison, the PAM would predict that a long-lag voiceless stop in Spanish would assimilate to the corresponding English voiceless stop category regardless of word position due to phonetic similarity. This would likely mean that learners would not associate the lengthened VOT with anything other than the stop, at least initially. However, the SLM would predict that, because VOT duration for voiceless stops is position-dependent in English,

the L2 learner of Spanish would not associate long-lag VOT in Spanish in word-internal onset position to the English stop allophone in that same word position. Consequently, the SLM predicts that, with exposure to the long VOT cue for /s/-aspiration in Andalusian Spanish over time, a learner would eventually create a new phonetic category of [t^h] in Spanish word-internal onset position. Initially, though, it would be hypothesized that a learner might perceive a phonetic difference between the Spanish and English stops in word-medial onset position while not yet having created a new category. In this case, the learner would be expected to simply categorize it as a stop until receiving enough input to create the new category for an aspirated stop. Then, the aspirated stop would have to be categorized as an acoustic cue that signals the absence of or aspiration of /s/.

When these phonetic cues are taken together, the following predictions can be made regarding how learners, upon receiving sufficient input of each type of production, would map the L2 phones (Table 8). Of the productions listed in Table 8, the least likely to occur is the second, breathy phonation followed by short-lag VOT. This is because recent studies have found that short aspiration preceding a voiceless stop typically coincides with long-lag VOT, as in the third and fourth scenarios below, while longer, strident aspiration tends to coincide with short-lag VOT (Parrell, 2012). However, this is not always the case (Ruch & Harrington, 2014). To know for sure, future studies should use well-controlled mapping tasks with phonetic continua similar to Ruch and Harrington's (2014) L1 perception study to determine how words with each of these realizations of /s/-aspiration are mapped. Importantly, the current study focuses on the third and fourth cases below in which long-lag VOT is present and aspiration preceding the stop is either non-existent or is of the breathy phonation type. It is left to future studies to compare the perception of different cues and cue combinations to determine how learners weight these acoustic cues.

Table 8. Summary of PAM and SLM predictions for the perception of /s/-aspiration

Aspiration	VOT	PAM prediction	SLM prediction
Strident, voiceless ([h])	short-lag	[h] mapped to English /h/ category.	Not mapped to English /s/ or /h/ category. New phonetic category formed for coda [h] over time.
Breathy phonation	short-lag	Breathy phonation potentially mapped as part of the vowel. /CV.CV/ structure perceived.	Phonation probably not perceived as part of the vowel. New phonetic category formation for a lengthened vowel with breathy voicing eventually?
Breathy phonation	Long-lag	Phonation likely perceived as part of the vowel, VOT mapped to English stop category, inhibiting /CVs.CV/ categorization.	Phonation not perceived as part of the vowel and may become a new category. New category formed for post-aspirated stop in word-medial onset position over time.
No glottal activity	Long-lag	VOT mapped to English stop category initially, /CV.CV/ perceived.	New phonetic category must be formed for a post-aspirated stop with sufficient input.

Since the current study not only concerned with how learners will map /s/-aspiration orthographically, but also the speed and accuracy with which they access words from the mental lexicon when /s/-aspiration is present, it is also important to hypothesize about learners' performance on a lexical decision task. The first prediction is that, since learners are expected to perceive the absence of a coda consonant prior to extensive exposure to /s/-aspiration in Andalusian Spanish, they will be significantly less accurate and respond more slowly to words that contain /s/-aspiration than words that contain a coda sibilant. In other words, they will tend to consider words containing /s/-aspiration to be non-words initially and will treat them as such in responding on a lexical decision task. It is possible that learners initially respond accurately to some of the words that contain /s/-aspiration because lexical items have other phonemes and acoustic cues that will help lexical processing. However, the overall pattern is not hypothesized to show a high degree of accuracy initially.

Over time, then, learners who are exposed to /s/-aspiration in Andalusian Spanish during study abroad are hypothesized to significantly reduce their response time and increase their accuracy when responding to words that contain /s/-aspiration. This is because these learners will have input that leads to a restructuring of the L2 phonological system so that lexical and phonological representations are updated to include /s/-aspiration as a legitimate allophonic variant of the phoneme /s/. Learners who are not exposed to Andalusian Spanish, on the other hand, are hypothesized to not show this pattern and should treat words containing /s/-aspiration in the same manner (i.e., as non-words) when tested at two different time points.

Models of Spoken Word Recognition

The models described above are useful for determining how L2 phones might be perceived according to L1 categories and how categorization may change over time, but they are also limited in scope for two main reasons. The first is that they focus on the segmental level of perception and do not describe the process of going from the perception of an acoustic signal to the recognition of a word. The second reason is that they have focused on the perception of phonemic contrasts and have not sufficiently explained how phonological variation (i.e., allophonic variation) should be processed by L2 learners. Table 8 shows that it is quite difficult using the current models to hypothesize about L2 learners will categorize Andalusian /s/-aspiration due to the presence of multiple cues that vary significantly from production to production and among different speakers and in different contexts. What is important here is not whether [s] and an aspirated variant and a post-aspirated [t^h] and an English /t/ can be discriminated, which is the focus of the SLM and PAM. Discrimination should be very good given that they would map to two different L1 categories ([s] → English /s/, [t^h] → English /t/) according to the PAM. What is important to the current study is whether L2 learners map aspirated variants of /s/ to the phonemic category of /s/ and also how lexical access is affected by /s/-aspiration. If

an aspirated variant is not categorized as an allophone of /s/, word recognition will be negatively affected. For this reason it is important to look at both identification *and* lexical access.

Over the years, various models of word recognition have been proposed and have attempted to explain how phonological variation is processed by the perceptual system in the recognition of words (i.e., lexical access). Models of word recognition agree that there are two fundamental processes to word recognition: activation of competitors and selection. The general view on the basic mechanisms allowing word recognition is described in the Cohort model of Marslen-Wilson (1987), according to which there is a process of activation of candidates based on the acoustic signal and how well the signal matches with lexical entries. These candidates compete in parallel. As the speech signal unfolds, any phonetic mismatch in the input will de-activate candidates one by one, until the point of recognition is reached, that is when there is only one candidate left (the winner) and all others are deactivated. Studies on activation have generally found that multiple word forms that are similar to the input are activated simultaneously. According to Luce & McLennan (2005) the evidence shows that consistencies between the input and mental representation at any place in a word, not just at the onset of a word, can result in activation. It would be predicted, then, that when L2 learners are confronted with /s/-aspiration in word-medial coda position, the input and mental representations for the word should be mismatched if the learner does not have experience with /s/-aspiration in the input. A word may be activated initially based on the beginning of the word, but then the aspirated variant would cause deactivation, or at least slow lexical activation. With increasing experience with /s/-aspiration, lexical activation should get faster and more accurate as an aspirated /s/ is categorized as an allophone of /s/ in the L2 phonological inventory.

Some of the proposed models have taken a bottom-up approach to perceptual processing that begins with phonetic or feature-based representation (allophonic level), proceeds to categorization into abstract phonemic units, and ends with combining phonemes and semantic information to process a

word (e.g., *TRACE*, McClelland & Elman, 1986; *Shortlist*, Norris, 1994). While some of these models assume sublexical levels of representation (e.g., abstract phonemic units), others do not (e.g., Distributed Cohort Model, Gaskell & Marslen-Wilson, 1997, 1999, 2002). The DCM bypasses sub-lexical (i.e., phonemic) processing and proposes that the acoustic-phonetic input is mapped directly onto form-based lexical representations. All of the models, however, “posit coding an acoustic-phonetic signal into either abstract phonetic features (in *TRACE* or DCM) or phonemes (in *TRACE* and *Shortlist*) that vary neither as a function of time, rate, phonological context, or talker” (Luce & McLennan, 2005, p. 594). Interestingly, though, recent research has found substantial evidence that surface details such as talker information and physical detail of the acoustic stimulus are not discarded during processing (see Luce & McLennan, 2005 and Pisoni & Levi, 2007 for reviews). According to Luce and McLennan (2005)’s review of the literature there is significant evidence to support detailed, exemplar-based representations in addition to abstract representations. An exemplar-based processing system is adaptive and exemplars change over time with exposure to new input. However, and very importantly, McLennan, Luce, and Charles-Luce (2003) and Pisoni and Levi (2007) come to the same conclusion, based on many research studies, that *both* types of representation (i.e., abstract and exemplar) must be incorporated into a theory of word recognition and processing depending on the speed of processing. In other words, *both* types of lexical representation can occur depending on task and processing time.

This recent line research on word recognition indicates that learners will store and use acoustic-phonetic and indexical information in the processing of allophonic variation related to /s/-aspiration during a study abroad experience. Schmidt (2011) suggested that a hybrid model that includes the indexing of contextual and social information was likely best suited for interpreting the results of her study of the perception of Venezuelan and Argentinian /s/-aspiration by learners of Spanish. She argued that such a model could explain how learners improve perception as they gain experience, given that exemplar models hold that listeners store instance-specific information about speech and that the

mental representation of speech consists of clusters of all experienced tokens that are indexed alongside contextual and social information (Bybee, 2007). Therefore, L2 learners will accumulate all instances of /s/-aspiration that they experience over time alongside contextual information and therefore increase their ability to correctly access words containing /s/-aspiration from the mental lexicon more quickly, as well as understand the social and contextual implications of the production of /s/-aspiration. With these considerations in mind the current study will measure the perception of Andalusian /s/-aspiration using two task paradigms: a forced-choice identification task and a lexical decision task, which have both been shown to be useful in the investigation of the acquisition of phonological variation.

Two Task Paradigms for Measuring the Perception of L2 Phones

Forced-choice Identification

In studying how L2 learners identify non-native speech sounds, a very common method has been the forced-choice identification task. This task has taken various forms, but with similar designs. Typically the participant will hear an aural stimulus and be required to identify, through written or computer-based means, a target orthographic or phonetic symbol (Carlet & Cebrian, 2014; Escudero & Williams, 2011, 2012; Flege & Eefting, 1987; Guion, Flege, Akahane-Yamada, & Pruitt, 2000; Hamann, Boersma, & Cavar, 2010; Harnsberger, 2001; MacKain, Best, & Strange, 1981; Rose, 2012), word (Aliaga-García, Mora, & Cerviño-Povedano, 2011; Flege, 1984; Flege & Hillenbrand, 1986; Motohashi-Saigo & Hardison, 2009; Ingram & Park, 1997; Schmidt, 2011; Shoemaker, 2010; Sonu et al., 2011), phrase (Tajima, Port, & Dalby, 1997), tone (Wang, Jongman, & Sereno, 2003), or an association with a similar keyword that represents the target sound in the aural stimulus (Flege, Bohn, & Jang, 1997; Flege, MacKay, & Meador, 1999; Mayr & Escudero, 2010). Sometimes goodness-of-fit ratings accompany the orthographic mapping based on the Perceptual Assimilation Model's (Best, 1995) assertion that L2

sounds are mapped onto L1 categories on a scale between being a poor exemplar or a good exemplar of the L1 category (e.g., Rose, 2012). This rating can help to determine whether the L2 sound will be considered part of an existing L1 category, or whether a new sound category for the L2 sound will be created. It also helps in determining how difficult a sound contrast will be to discriminate.

Research in psycholinguistics has shown that bilinguals cannot avoid the activation of lexical information in both languages even upon hearing just one of the languages (Brysbaert, 1998; Dijkstra & Van Heuven, 1998; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Marian & Spivey, 2003; Van Heuven, Dijkstra, & Grainger, 1998). The question that follows is whether competing orthography-phonology mappings in the two languages, or even within one language, interfere in the identification of phones. For example, Escudero and Wanrooij (2010) found that L1 orthography both positively and negatively affected Spanish listeners' categorization of Dutch vowels depending on the vowel contrast in question. Schwartz, Kroll, & Diaz (2007) discuss that evidence has been found that identifying a word visually is affected by the reliability of mappings between orthographic and phonological codes (p. 107). They review studies that have found that when an orthographic code maps onto multiple phonological codes or a phonological code maps onto multiple orthographic codes performance can be hindered. Studying whether bilingual processing of lexical items (L1 English, L2 Spanish) is affected by the reliability of orthography-phonology mappings across languages, they found that "degree of consistency between orthographic and phonological codes influences the manner in which cross-language competition is manifest" (p. 108). Therefore, an important question that must be answered regarding the forced-choice identification paradigm is how reliable the orthography-phonology mappings are between the L1 and L2 of the learners for the target sound(s) under investigation and what the task is actually able to tell us about acquisition.

One issue with a forced-choice identification task as described above is that orthographic representations do not always have a one-to-one correspondence with the phonology. For example, in

Spanish, orthographic <s> only corresponds to one phoneme in Spanish: /s/. The same is true in English. However, the Spanish grapheme <s> can represent multiple allophones of the phoneme /s/ depending on the context. Before a voiced consonant, <s> is generally, but not always, produced as the voiced alveolar fricative [z] (Schmidt & Willis, 2011), while before a voiceless consonant, it is usually produced as voiceless [s]. In terms of sociolinguistic variation, orthographic <s> in some dialects can also be the orthographic representation in a word that is pronounced with an aspirated [h] or a deleted [s], or in the case of Andalusian Spanish, <s> can also represent the combination of multiple acoustic cues that define aspirated /s/ (O'Neill, 2010; Ruch, 2013; Ruch & Harrington, 2014; Torreira, 2006, 2007, 2012). The opposite can also be true, that multiple orthographic codes can map onto one phonological code. This can be seen in the Castilian theta /θ/, which can be represented orthographically by <ci>, <ce>, <z>, and in some dialects, <s>. In both English and Spanish, /s/ can be realized orthographically as <c> or <s> depending on the following phonological context. In this way, English and Spanish are similar, though English /s/ does not have the allophones [h] (or other aspirated variants) or [z] as it does in Spanish, since /h/ and /z/ are phonemic in English. We should expect, then, that the identification of the voiceless alveolar fricative [s] as orthographic <s> in Spanish should not be difficult for native English speakers. The more pressing question is how native English speakers learning Spanish as a second language identify aspirated (i.e., [h] or [p^h t^h k^h]) variants of /s/ orthographically, and what that tells us about their L2 phonological system.

A good place to look to answer this question is Schmidt's (2011) dissertation, which tested the perception of /s/-aspiration (i.e., [h] in Argentina and Venezuela) by L2 learners and native speakers using a forced-choice identification task. The results for the orthographic identification of the [h] variant of /s/ patterned as would be expected based on the experience of the participants. There was one native speaker group from a dialect that aspirated /s/ and another from a non-aspirating dialect. The native speakers from the /s/-aspirating region were significantly more accurate at labeling [h] as <s>

than NSs from the non-aspirating region. Each group had a small number of participants who performed more like the other group, usually due to exposure through personal contact with speakers who aspirate /s/. Schmidt (2011) also tested various groups of English-speaking learners of Spanish with different levels of experience with the target language and different experiences studying abroad including differences in location of study abroad in dialects that exhibit /s/-aspiration and others that did not. The results showed a clear pattern that differentiated the groups who had received exposure to aspirated /s/ and those who had not. The more advanced learners began to identify [h] as orthographic <s>, whereas the beginner learners did not, tending to label non-words with coda [h] as having *no coda*. This pattern is what would be expected based on the experience of the learners and the native speakers' places of origin. Importantly, this result came about from the use of non-words in the task. This is important because the use of non-words shows a pattern that cannot result from word knowledge. In fact, what Schmidt (2011) showed was that those with the most experience with /s/-aspiration, upon hearing a non-word with coda /s/-aspiration, more consistently applied the phonotactic constraints on /s/-aspiration to the non-words. This indicates that, even without word knowledge, phonological patterns were applied, signifying that /s/-aspiration was a part of the L2 phonological system to some degree for each of those participants. It shows that by using the orthographic label <s> they were able to show that they had categorized /s/-aspiration as a variant of the phoneme /s/ in Spanish in word-internal coda position since <s> corresponds to the phoneme /s/. For the L2 learners it signified that some had restructured their interlanguage phonology while others had not.

In sum, even though orthographic representation is not always directly related to phonological representation, the results of Schmidt (2011) indicate that the forced-choice identification task will be able to show how learners categorize /s/-aspiration in Andalusian Spanish. This categorization is at least a connection between the acoustic phonetics of the aspiration and the orthography, but is also likely indicative of deeper phonological restructuring.

Lexical Decision

Another important question to answer regarding the acquisition of phonological variation is how much phonological variation inhibits learners from accurately accessing spoken words from the mental lexicon. This is important in a very practical sense because if a word is not accessed or is incorrectly accessed, comprehension and thus communication could be impeded. In order to determine whether Andalusian /s/-aspiration hinders lexical access and whether there is change over time, a lexical decision task will be used. The primary purpose of lexical decision tasks is to measure how accurately and quickly (i.e., response time) participants classify aural stimuli as real words or non-words in the target language. This measure can be used in a variety of ways (e.g., no priming, form priming, semantic priming, long-term repetition priming). This type of task recently been used to measure the perceptual processing of variable phonological input in L1 Spanish (Boomershine, 2006) and other languages such as English (Sumner & Samuel, 2009). In this type of experiment, results that show that the participants can accurately and quickly respond to real words that do not contain dialect-specific phonological variants as well as accurately rejecting non-words both with and without phonological variation, while at the same time having difficulty responding accurately and quickly to real words that do contain dialect-specific phonological variants, indicate that the phonological variation causes processing difficulty due to the absence of the variant in the phonological representation of a word.

Boomershine (2006) studied the effects of linguistic experience on native speakers' ability to identify and process dialect-specific phonological variation. She tested the perception Puerto Rican Spanish /r/-lateralization, /s/-aspiration and deletion, and /n/-velarization with 33 participants from Puerto Rico (n=13) and Morelos, Mexico (n=20). One of the tasks she used was a lexical decision task, which tested the effect of a speaker's dialect on their ability to decide if a word produced by someone from their own dialect or a different dialect was a real or non-word. She was able to show that there was a significant difference in the response times based on the phonological variable in question and the

dialect of the speaker. She showed that Mexican listeners had more difficulty than Puerto Rican listeners with coda /s/-aspiration because they were not used to /s/-aspiration in their dialect. Boomershine (2006) also noted that her identification task, in which the participants identified the dialect of different speakers, supported the results of the lexical decision task. The phoneme /s/ and its variants were clearly what listeners used the most to determine what dialect the speaker was from, and the greatest difference in lexical decision accuracy between dialect groups was based on the productions of this phoneme.

Sumner and Samuel (2009), although not a study about Spanish, is an important study that made use of three types of lexical decision tasks to study the perception of phonological variation in American English by L1 English speakers. The variable was the production of New York City /r/ and General American /r/ by three groups of participants: NYC speakers who use NYC /r/ in production, NYC speakers who do not use NYC /r/ in production and non-NYC speakers from another dialect region that use General American /r/ and not NYC /r/ (GA group). The three tasks were all variations on priming experiments using the LD paradigm. The first was a form priming task that examined the effects of surface features on immediate processing. Listeners heard a prime followed by a phonologically related target. An identical prime usually leads to faster response times than when the target is preceded by a phonologically unrelated prime. They note that this paradigm allowed them to determine whether “phonologically deviating from a dialect variant has an effect on processing that variant” (p. 490). This experiment showed that those with previous exposure to NYC /r/ (the two NYC groups) were similar in their priming effects, while the General American /r/ group was significantly slower in their response times. This group exhibited faster response times only when the target was from their own dialect. The GA group also showed higher error rates when responding to NYC targets than the other two groups. They concluded that “listeners who have experience with both dialects are more flexible in form processing: these listeners are flexible in that they do not have a rigid criterion for a particular form

when it is a regular variant” (p. 493). Over time, they argue, their exposure to both variants makes them more comparable perceptually. The GA group was only exposed to one variant.

The second task was a semantic priming lexical decision task. This task was used to study the effect of different variants on lexical access as a function of experience. Given that mismatching information often hampers lexical access (Connine, Blasko & Titone, 1993; Gaskell & Marslen-Wilson, 1993; Marslen-Wilson, 1987; Soto-Faraco, Sebastian-Gallés, & Cutler, 2001; Sumner & Samuel, 2005), they “used canonical r-ful primes and semantically-related targets as a baseline for all participants...[and they] assessed how listeners processed the phonological variants by comparing the effectiveness of r-less primes in facilitating the processing of semantically-related targets to the canonical baseline priming” (p. 493). They hypothesized that NYC r-less words should not prime semantically related targets unless r-ful/r-less variation is a part of their phonology. Indeed, that is what they found. The two NYC-exposed groups exhibited priming effects for r-less and r-ful words primed by semantically related targets, while the GA group only showed priming for r-ful words.

In order to see how the variants were represented and stored in the long-term, they used a long-term repetition priming LD task. This task separated primes and targets by 20 to 30 minutes in two different blocks. If the variants are stored for the long term, primed words should be accessed faster than unprimed words in spite of the length of time between them. In this task, they found that the two NYC groups, though equivalent on the first two tasks, were actually different, showing only the NYC group that produced r-less variants as having these forms represented in the long term, while the NYC group that only produced r-ful forms did not. They termed this second group “fluent listeners,” as a sort of intermediate status between the NYC r-less group and the GA group as a result of their experience. Additionally, the GA group was slower overall because of not having r-less forms in their dialect. The r-less primes produced poorer facilitation for them than r-ful primes.

Sumner and Samuel (2009) is an important study because it used the LD paradigm to understand the processing of dialect-specific phonological variation. They were also able to tease apart groups based on their experience with the phonological variants, even showing differences between the two NYC groups based on their different degrees of use of the variants. Their tasks showed processing deficits for those not exposed to or not producing the variants. Sumner and Samuel (2009)'s study is very relevant to the current study because, although it was a study on the perception of phonological variation within the L1 of the participants, it is one of the few studies to investigate the perception of dialect-specific phonological variation in view of one's exposure to the variants, particularly when using the LD paradigm. The current study is similar in its goals to Sumner and Samuel (2009) in that, at least on the lexical side, it seeks to determine whether learners who gain intensive exposure during study abroad gain improved lexical access when confronted with words containing /s/-aspiration. In other words, when exposed to /s/-weakening variation, learners should increase their speed and accuracy of lexical retrieval of words that contain /s/-aspiration, indicating that they are acquiring the ability to encode /s/-aspiration in their phonological and lexical representations.

Summary of the Motivations for the Current Study

To summarize, the review of research on context of learning in SLA research, the acquisition of variation in a second language, the acquisition of dialect-specific variants, and L2 phonology has led to the following important contributions and motivations for the current study. First, we have learned that variable linguistic forms are acquired late in the SLA process (Geeslin, 2011b). Sometimes learners reach native-like use (rates of use and/or predictors of use) and other times they do not. At times, the path of development is linear, but other times it is not and learners can experience gains and losses before restructuring the linguistic system to accommodate variable forms. We have seen that various factors can contribute to the acquisition of variable forms, such as study abroad experience, dialect of exposure,

attitudes toward the target variants, language proficiency, and years of study of the L2. However, there are still important questions that remain unanswered.

Research on the perception of dialect-specific L2 phonological variants is scant and it is unclear how it may be similar to or different than findings for variable morphosyntax and the production aspect of dialect-specific L2 phonological variants. The recent findings of Schmidt (2011) indicate that the path to perceiving dialect-specific variants may be similar to the acquisition of morphosyntactic variation in terms of being acquired late in the SLA process and being positively affected by exposure to the dialect in which the variation occurs. Perception research is also necessary because studies have shown that learners tend to not produce variable phones. However, this does not necessarily mean that their L2 phonological system is not categorizing phonological variants and leading to increased comprehension. Studies on the perception of dialect-specific variants will also help us understand how individual factors are related to speech perception and how this may be different than for speech production. For example, grammatical proficiency may not condition the production of dialect-specific variants given the choice that is involved in producing them and the influence of seemingly stronger factors like social networks, while the perception of dialect-specific variants may be affected differently by these factors and proceed subconsciously despite differences in grammatical proficiency or other factors that differentiate learners. Finally, research on the acquisition of variable forms in the study abroad context has not yet investigated the development of the perception of dialect-specific variants longitudinally. This is important because it could show the restructuring of the linguistic system over time, which would provide clues as to how it occurs and what factors predict such restructuring. Thus, research on context of learning in SLA is an important avenue that must be explored for speech perception.

The research that has been conducted to date on context of learning has found that study abroad is most beneficial for the broader aspects of communicative competence, such as fluency and communication strategies, rather than the minutiae of morphosyntax and phonology (Freed, 1995;

Mora, 2014; Valls-Ferrer & Mora, 2014). The study abroad context is meaning-oriented and learners tend to focus more on the aspects of language that will help communicate, while the form-oriented classroom context tends to help learners acquire specific morphosyntactic forms and explicit phonetics instruction can help learners' pronunciation of L2 phones (Aliaga-Garcia & Mora, 2009; Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999; Isabelli-Garcia, 2010; Pisoni & Lively, 1995). Importantly, though, the study abroad context *has* often been shown to be beneficial for the acquisition of sociolinguistic competence, or the aspects of the target language for which native speakers show variation (Geeslin et al., 2010, 2012; Regan, 1995, 1996; Regan, Howard & Lemée, 2009; Sax, 2003; Thomas, 2004). Furthermore, the location of study abroad (i.e., dialect of exposure) appears to be important in that learners have tended to acquire the norms of the dialect to which they are exposed (e.g., Geeslin et al., 2012; Geeslin, Fafulas & Kanwit, 2011; Kanwit & Solon, 2013; Salgado-Robles, 2011, 2014). However, not all studies have found exposure to dialectal variation to lead to the use of variable forms by learners (Fox & McGory, 2007; Geeslin & Gudmestad, 2008b; George, 2014; Knouse, 2012; Ringer-Hilfinger, 2012). Potential explanations lie in the differences between individuals such as grammatical proficiency, motivation, attitudes toward the dialect or linguistic feature in question, and factors such as learners' social networks. Several of these studies showed that many students did not want to sound like speakers from the dialect in which they studied while others did (e.g., Ringer-Hilfinger, 2012). George (2014) found that social networks with native and non-native speakers played a role in the use of dialect-specific phonological variants, which is in accord with recent research on the importance of social network formation during study abroad and its effects on linguistic gains (Dewey et al., 2013; Martinsen et al., 2010).

An important consideration that has not been addressed in most study abroad studies due to their focus on production is the role that speech perception plays in the acquisition of variable L2 phonology over time during SA. Given that production studies have shown very low rates of use of

variable target phones, an important question is how learners perceive the phones. It very well could be that at least some of the learners do not produce the target phones simply because they cannot perceive them and categorize them accurately in the L2 phonological system. In other words, it is possible that the variation presented by the target phones hinders lexical access and/or creates confusion when they map them onto their L2 phonological inventory. Perhaps other factors such as native speaker contact and target language use, grammatical proficiency, or study abroad program play a role in the perception of dialectal variants longitudinally during SA. The current study investigates whether learners come to perceive a regional allophonic variant of the phoneme /s/ in word-internal coda position in Western Andalusia (Seville) over time. The study focuses on how they perceive the acoustic cues of the variant (i.e., categorization and lexical access) and how the individual factors of exposure (i.e., study abroad vs. AH) NS contact, target language use, grammatical proficiency and study abroad program predict their perceptual patterns. This study makes a contribution to the field by adding to the yet very small body of research on the acquisition of dialect-specific phonological variation, for which the learning context is potentially very important due to increased opportunities for input that can be used in different ways by learners. Thus, the current study is guided by the following research questions.

Research Questions

1. How is word-internal coda /s/ aspiration in Western Andalusian Spanish perceived by L2 learners?
 - a. Do L2 learners associate aspirated variants of coda /s/ with orthographic <s>? What are the patterns of association when aspirated variants are not identified as <s>?
 - b. Does the presence of an aspirated variant of coda /s/ hinder accuracy and/or response times for L2 learners' lexical access?
2. Does longitudinal exposure to Western Andalusian Spanish affect word-internal coda /s/ orthographic association and lexical access patterns (i.e., study abroad versus at-home)?
3. Do the following individual factors play a role in predicting orthographic association and lexical access patterns at the beginning and the end of one semester?
 - a. Grammatical proficiency
 - b. Contact with native speakers
 - c. Spanish language use by the learner
 - d. Study abroad program

Chapter 3: Methods

This chapter will describe the methods used in order to test the research questions. It will begin with a description of the participants, followed by an explanation of the instruments and procedures used for data collection, and will conclude with a description of the data analyses.

Participants

The participants in this study were native speakers of Spanish (N=45) from two different groups and English-speaking learners of Spanish as a second language (N=116). The two native speaker groups consisted of 35 native speakers of Spanish in Seville, Spain and 10 native speakers of Spanish from non-aspirating Spanish-speaking regions who were graduate students at Indiana University Bloomington. Of the 116 L2 learners who participated in the first data collection in late August and early September of 2014, 77 were enrolled in one of two study abroad programs in Seville, Spain (“Study Abroad 1” and “Study Abroad 2” groups) during the fall semester of 2014 and 39 L2 learners were enrolled in at least one Spanish course during the same semester at Indiana University (“At-home” group). 99 of the original 116 L2 participants at Time 1 also completed the second data collection in mid-November of 2014 (Time 2). After exclusions (to be described in detail below), 73 L2 learners of the original 116 were included in the analysis. Of the 35 native speakers in Spain, 33 were included in the final analysis. All 10 of the non-aspirating speakers were included in the analysis (“Non-aspirating” group). All participants who participated at Time 1 and Time 2 completed two speech perception tasks each time, including a forced-choice identification task and a lexical decision task. The L2 learners, both at Time 1 and Time 2, also completed a word familiarity questionnaire, a grammar test and a questionnaire both at the beginning and end of the semester. The native speaker groups were tested once and also completed the grammar test and one background questionnaire, but did not complete the word familiarity questionnaire. Each task and the procedures will be described in detail following a description of the participant groups.

Native Speaker Groups

The primary native speaker control group represents the target dialect of Western Andalusian Spanish, primarily Seville, and consisted of 35 NSs from Seville or nearby towns in Western Andalusia (henceforth the “SEVILLE” group). The second NS group (henceforth the “Non-Aspirating” group) consisted of 10 NSs of Spanish from Spanish-speaking regions in which coda [s] is maintained. The SEVILLE group completed the study in Seville, while the Non-Aspirating group completed the experiment on the Indiana University Bloomington campus. The recruitment of the SEVILLE group took place in Seville during two research trips to Seville in early September of 2014 and 11 weeks later in mid-November of 2014, and this group participated either in a computer lab at one of the study abroad program facilities, in the home of the participant or another person’s home, or in another location of mutual agreement. The Non-Aspirating group was recruited through personal contacts at Indiana University and this group participated in computer labs or a linguistics lab at the university.

SEVILLE group. Of the 35 NSs from the SEVILLE group that participated in the study in September or November of 2014, two had to be eliminated from the analysis. One had spent many years living in the United States during graduate school, much longer than any of the other NSs had lived outside of Western Andalusia. Another was eliminated due having been born in Venezuela and lived there for approximately half of his life before moving to Seville at the age of nine. Of the 33 remaining NS participants, 20 were male and 13 female and were between the ages of 18 and 41, with a mean age of 28.12 years (SD = 6.41, Median=28). According to the background questionnaire, all but three were originally from Seville. The participants had lived in Seville for more than 20 years (N=26), between 15 and 20 years (N=3), between six and ten years (N=1), between two and five years (N=2), and for one year or less (N=1). The three NSs who were not originally from Seville were from Huelva, Algeciras and Málaga, which are located in Western Andalusia, and they had lived in Seville for between one and ten years. Of the 33 participants in the SEVILLE group, most responded that they had lived their entire lives

in Western Andalusia (N=26), while the rest (N=7) had lived the majority of their lives in Western Andalusia but had lived in one to three other places for a period of time, typically for one year or less, with a few exceptions (Table 9). The most notable is SEVILLE09, whose mother is from the United States and whose father is from Seville. SEVILLE09 is 18 years old and has lived in Seville most of her life with the exception of ages two through five when she lived in Houston, Texas.

Table 9. Duration of stays outside of Western Andalusia (SEVILLE group)

Participant	Place	Duration
SEVILLE07	Cardiff, Wales, UK Bloomington, Indiana	1 year or less each
SEVILLE09	Houston, Texas	3 years
SEVILLE11	England	1 year or less
SEVILLE15	Manchester, England	1 year or less
SEVILLE18	Nijmegen, Netherlands Limerick, Ireland Oxford, England	1 year or less each
SEVILLE21	Buenos Aires, Argentina	1 year or less
SEVILLE29	Paris, France	1 year or less

Most of the SEVILLE participants reported knowing at least one language other than Spanish to some degree (N=29), the most common of which was English (Table 10). Other languages included French, Italian, German, Portuguese, Catalan, Arabic and Japanese. The language with the most proficient speakers was English, with 16 participants reporting advanced (N=12) or near-native (N=4) proficiency. French was the second most known language on the list and also had the second-highest level of proficiency among the participants.

Table 10. Self-reported knowledge of languages other than Spanish by the SEVILLE participants

Language	Self-reported level	# of participants
English	Beginner	1
	Intermediate	12
	Advanced	12
	Near Native	4
French	Beginner	6
	Intermediate	6
	Advanced	2
Italian	Beginner	7
	Intermediate	1
German	Beginner	3
	Intermediate	1
Portuguese	Beginner	1
	Intermediate	1
Catalan	Beginner	1
Arabic	Beginner	1
Japanese	Beginner	1

The participants were asked on the background questionnaire about the highest level of education they had achieved. Most reported either being enrolled in the university at the time of data collection (N=11) or that they had completed a bachelor's degree (N=12). Other education levels reported included high school (N=2), vocational training (N=4), or a master's degree (N=4). The participants reported having various careers, but most reported working in an academic environment. The jobs reported include professor, elementary school teacher, administrative assistant, corporate investigator, study abroad program director or administrator, study abroad program staff, study abroad program orientation guide, computer scientist, lifeguard, event coordinator, salesperson, X-ray technician and electrician. Finally, the SEVILLE group also completed the grammar test as a baseline against which to compare the L2 groups' grammar scores. The mean score on the grammar test for the SEVILLE group was 19.24 (SD = 0.99, Range = 16 - 20) out of a possible 20 points.

Non-Aspirating group. Of the 10 native speakers of Spanish who were from regions of the Spanish-speaking world in which /s/ is maintained, six were male and four were female. The age range

was between 30 and 43 years old with a mean age of 35.4 years (SD = 3.95, Median=36). According to the background questionnaire, the 10 participants reported being from Spain (Burgos, Ciudad Real, Palencia), Mexico (Durango, Ciudad Juárez, Toluca), Ecuador (Quito), Peru (Huancavelica), and Colombia (Bucaramanga, Bogotá).

Table 11. Demographic information for Non-Aspirating group.

Participant	Place of origin	Other places lived	Duration	Occupation
NON-ASP01 F, 36	Burgos, Spain	Soria, Spain Valladolid, Spain Morgantown, WV Athens, OH Bloomington, IN	2-5 years 1 year or less 2-5 years 2-5 years 6-10 years	Spanish lecturer
NON-ASP02 M,32	Huancavelica, Peru	New Jersey, USA Bloomington, IN	6-10 years 6-10 years	Teacher
NON-ASP03 F, 33	Durango, Mexico	Chicago, IL Normal, IL Tremont, IL Bloomington, IN	6-10 years 2-5 years 2-5 years 2-5 years	Spanish instructor
NON-ASP04 F, 38	Ciudad Real, Spain	Ohio, USA Limoges, France Florida, USA Missouri, USA Bloomington, IN	2-5 years 1 year or less 1 year or less 6-10 years 1 year or less	Spanish lecturer
NON-ASP05 M, 37	Bucaramanga, Colombia	USA (unspecified)	10-15 years	Spanish lecturer
NON-ASP06 M, 43	Ciudad Juárez, Mexico	Juárez Monterrey, MX Indiana, USA Ohio, USA Connecticut, USA	10+ years 10+ years 2-5 years 2-5 years 1 year or less	Spanish instructor
NON-ASP07 M, 30	Quito, Ecuador	Bloomington, IN	2-5 years	Research Assistant
NON-ASP08 M, 31	Palencia, Spain	Bergen, Norway Bloomington, IN	1 year or less 6-10 years	Librarian
NON-ASP09 M, 36	Bogotá, Colombia	Bath, England Bloomington, IN	1 year or less 2-5 years	Graduate Student
NON-ASP10 F, 38	Toluca, Mexico	Federal District, MX Bloomington, IN	10+ years 2-5 years	Unspecified

All but one of the Non-Aspirating participants reported knowledge of other languages. These languages included English, French, Italian, German, Portuguese and Catalan. The number of participants with each self-reported proficiency level for each language is summarized in Table 12.

Table 12. Self-reported knowledge of languages other than Spanish by the Non-Aspirating participants

Language	Self-reported level	# of participants
English	Advanced	3
	Near Native	5
French	Beginner	3
	Advanced	1
Italian	Advanced	1
German	Beginner	1
Portuguese	Intermediate	2
Catalan	Beginner	1

Regarding the Non-Aspirating group's education level, all 10 participants reported having completed a master's degree (N=7) or a doctoral degree (N=3). Almost all of the participants were graduate students at the time of data collection and were employed as Spanish instructors in the university. One participant reported being a librarian (NON-ASP08) and another did not specify an occupation (NON-ASP10). Finally, the mean score on the grammar test for the Non-Aspirating group was 19.4 (SD = 0.80, Range = 18 - 20) out of a possible 20 points.

L2 learner groups

The L2 learner groups were L1 English-speaking learners of Spanish in two primary groups. The study abroad group consisted of 77 learners at Time 1 (i.e., September, 2014) who studied abroad for one semester in Seville Spain during the fall semester of 2014. The study abroad group was recruited via two study abroad programs in Seville. All of the participants from the study abroad group were enrolled in intermediate or advanced courses during the semester abroad. Recruitment took place during orientation programs and in some cases through the visitation of classes by the researcher. The At-home L2 group consisted of 39 learners at Time 1 who were enrolled in at least one second- or third-year

Spanish course at Indiana University during the same semester. These courses were chosen because it was important to control for the prior study abroad experience of the At-home group to ensure that none had been abroad for a lengthy period of time (i.e., more than one month) prior to data collection. This was controlled by asking the learners about their travel history to Spanish-speaking countries for three weeks or more at one time on a background questionnaire. Typically, learners that are enrolled in second- or third- year courses have not yet studied abroad, since the most common years for study abroad are the junior and senior years (Institute of International Education, 2014). Learners enrolled in Spanish courses at the beginner level were excluded from the study given that one of the tasks requires more knowledge of Spanish vocabulary than a typical beginner student would likely have.

Study abroad program #1 (henceforth "Study Abroad 1"). The study abroad participants were enrolled in one of two study abroad programs in Seville. The first study abroad program is small and draws students from a small number of mostly private colleges in the United States. This program, Study Abroad 1, employs an on-site program director, who is from the United States but is married to a native of Seville, two receptionists who are from Seville, four professors, and one student affairs coordinator, who is from the U.S. Of the four professors, two are from Seville or a nearby part of Western Andalusia, while the other two are from Valladolid and the Basque country but have lived in Seville and taught at Study Abroad 1 since the mid-1980s. In order to verify whether the four professors aspirated /s/ in class, the researcher observed one class taught by each of the four professors and it was determined that none of them produced /s/ aspiration during that particular class period. However, it was one of the first days of class of the semester and it is possible that as the students became more comfortable the professors' speech became less careful. However, no further observations were conducted.

Before the students' arrival in Seville and also at the in-country orientation, Study Abroad 1 administered two placement tests to the students in order to determine whether they would enroll in beginner, intermediate, or advanced classes. The first test was completed before arriving in Seville and

was the WebCAPE Spanish placement exam. The second was an exam designed by the faculty at Study Abroad 1 and was taken during the in-country orientation on the students' first full day in Seville. After the tests were scored, each student met with multiple faculty members to discuss the results, the most appropriate course level, and to enroll in the courses they would take for the entire semester. Of the 45 students enrolled at Study Abroad 1 during the fall semester of 2014, none tested into the beginner level. Three students were enrolled in intermediate courses, which are similar to traditional intermediate grammar courses in the U.S. The remaining 42 students enrolled in advanced courses that were more specific in content and included courses on different aspects of Spanish literature, history, culture, and advanced grammar. The courses were all taught in Spanish by the four professors and the program required students to sign a language pledge that they would only speak Spanish when inside the school. It was observed that, during school hours, the students generally abided by the pledge. However, when the researcher was in the school during evening hours students were observed speaking English frequently.

Of the 45 students enrolled at Study Abroad 1, 32 participated voluntarily at Time 1 after being recruited during orientation programming on the second full day after the students' arrival in Seville. At Time 2, 31 participants returned in order to complete the study. However, eight participants apart from the one who did not complete the study were excluded for various reasons. One participant was a native speaker of Korean. One participant reported having a history of hearing problems and also had extensive travel experience to Spanish-speaking countries. One participant reported having a friend from Seville with whom she had daily contact during the two years prior to studying abroad in Seville. And the five remaining eliminated participants had all spent more than one month in a Spanish-speaking country for various reasons, including mission trips, working as an interpreter for short term mission teams, vacation, nannyng or teaching. The countries they had visited included Bolivia, the Dominican

Republic, Peru, Ecuador, Mexico and Spain. One eliminated participant had spent two months in Madrid just prior to arriving in Seville.

After the elimination of the nine participants, the remaining 23 participants were all L1 speakers of English, did not speak any other L1s and did not have any immediate family members who spoke other L1s apart from English. Of the 23 participants, 20 were female and three were male. There were 20 students who were majoring in Spanish at their home university and three who were minoring in Spanish. One participant was enrolled in intermediate grammar courses at Study Abroad 1 and the other 22 were enrolled in advanced topic courses. The mean grammar test score at Time 1 for the Study Abroad 1 group was 13.07 (SD = 3.73, Range = 5 - 19) out of a possible 20 points.

With regards to the participants' Spanish education prior to studying abroad, approximately one half of the participants had taken Spanish in elementary and middle school while the other half had not. All participants had taken some high school Spanish with the majority having taken three or more years (N=18). One participant reported having attended a Spanish immersion school for six years in elementary school. In terms of the highest level of university Spanish courses that the participants had taken prior to study abroad, 11 had taken up to one or more second year courses, nine had taken up to one or more third year courses and three had taken up to one or more fourth year courses. In Table 13, those who had taken second year courses are listed as one to two years of university Spanish and those who had taken through third or fourth year courses are listed as having three or more years of university Spanish⁴.

⁴ Note that the responses for elementary, middle, and high school came from a different question on the questionnaire than the question regarding university coursework, which produced a different data output format. Thus, the university data has been synthesized to fit into the same table as the other levels of education. It is possible that some participants have not actually taken 3 or more years of college Spanish courses since they may have tested into a high level upon entering the university. For the purposes of the table, taking upper level coursework is considered the same as having 3+ years of university-level Spanish.

Table 13. Number of participants with each level of Spanish education prior to data collection (Study Abroad 1 group)

Education Level	0 years (# of participants)	< 1 year (# of partic.)	1-2 years (# of partic.)	3+ years (# of partic.)	Total #
Elementary school	12	6	1	4	23
Middle school	11	4	4	4	23
High school	0	1	4	18	23
University	0	0	11	12	23

On the background questionnaire, the participants were asked about their experience visiting Spanish-speaking countries for three weeks or more at a time prior to studying abroad in Seville. Most reported that they had never visited a Spanish-speaking country for three weeks or more (N=21), while two participants had. However, their stays were not longer than one month in duration and they were thus not excluded from the analysis. One participant visited Central America (i.e., San Salvador, El Salvador; Antigua, Guatemala; San Jose, Montenegro, and Arenal, Costa Rica) for 18 days for a mission trip (10 days) and family vacation (8 days). The other participant spent approximately one month in San Jose, Mexico for vacation.

The language background questionnaire also elicited information on the participants' contact with native speakers of Spanish during the two years prior to studying abroad in Seville, including information on the frequency of contact (daily, weekly, monthly, annually) and the type of relationship (e.g., employer, teacher, friend, acquaintance, co-worker, significant other). Of the 23 participants, 11 reported having had contact with at least one NS and seven reported daily or weekly contact with NSs. None reported contact with NSs from Spain, but did report contact with NSs from Mexico, Guatemala, Cuba, Nicaragua, Uruguay, Venezuela, Honduras, Ecuador, and Belize. The relationships between the participants and the NSs included friend (N=10), employer (N=1), teacher (N=1), significant other (N=2), acquaintance (N=3) and nanny (N=1).

Another item on the questionnaire elicited information on the participants' use of Spanish outside of the classroom during the year prior to the study abroad period. Table 14 displays the number

of participants that chose a given response. Only three participants reported speaking Spanish with native speakers on a weekly basis or more frequently, with the majority reporting very infrequent use of Spanish with native speakers. A higher number of participants reported speaking Spanish with non-native speakers, which is expected given that they were enrolled in Spanish courses. Most of the participants also reported infrequent participation in the use of various media in Spanish. However, listening to music in Spanish was the most frequent passive language activity in which the participants engaged in Spanish.

Table 14. Number of participants reporting Spanish use in each category prior to study abroad (Study Abroad 1 group)

Activity	A few times or less per year (# of participants)	1X/month (# of part.)	A few times per month (# of part.)	1X/week (# of part.)	A few or more times per week
Speaking Spanish with native speakers	14	2	4	2	1
Speaking Spanish with non-natives	7	0	3	3	10
Watching TV, videos, or movies in Spanish	13	7	0	3	0
Reading articles or books in Spanish	13	5	1	2	2
Listening to music in Spanish	7	3	4	2	7
Communicating online in Spanish	15	4	3	0	1
Texting in Spanish	16	3	3	0	1

Apart from studying Spanish, there were five participants who reported having studied other languages apart from Spanish and 18 who had not. One participant reported having an advanced level of proficiency in Latin. Another participant reported having studied Latin and Korean but at the beginner level. One participant reported intermediate proficiency in French. And one participant reported intermediate proficiency in American Sign Language. Lastly, one participant marked *yes* to having studied another language but did not list a language.

Study abroad program #2 (henceforth "Study Abroad 2"). The second study abroad program from which the participants in this study were recruited, henceforth referred to as the Study Abroad 2 group, draws from a larger number of universities and had many more students enrolled than Study Abroad 1. The program employs many professors and has a larger on-site staff than Study Abroad 1. The on-site staff consists of locals from Seville and two Americans. Some of the professors employed by Study Abroad 2 are also employed by local universities, teaching in both locations. However, some professors are exclusively employed by Study Abroad 2. Study Abroad 2 also employs a small group of local young adults whose job is to interact with students in Spanish between classes and take students on tours of the city in order to encourage the use of Spanish. The researcher observed students interacting in Spanish with these locals, however students were also observed speaking in English at the program site very frequently. Study Abroad 2 did not require students to sign a language pledge.

Study Abroad 2 has various sub-programs in which students enroll, including Liberal Arts, Advanced Liberal Arts, Business and Society, International Business and Culture, Communication, New Media and Journalism; Language and Society; and a Teaching Development Program. In all, 42 different courses were offered during the fall semester of 2014 among all of the programs combined. The students with the highest level of proficiency and those in the international business and culture program, took most or all of their classes at the University of Seville or the University of Pablo Olavide rather than in the classroom building owned and operated by the Study Abroad 2 program. Those who were enrolled in Liberal Arts, Communication, New Media and Journalism; or the Teaching Development Program, and also were not in the highest level of language proficiency took most of their classes at the Study Abroad 2 facility with other American students, with the possibility of taking one class at the university alongside native Spanish speaking students if they met the criterion for proficiency to be able to do so and chose to do so.

Before taking classes in their chosen curriculum, the Liberal Arts, Advanced Liberal Arts, Communication, New Media and Journalism; and Teaching Development Program students took a two-week intensive Spanish course taught by native Spanish speaking professors. The total number of students enrolled in the intensive session was 92. For the intensive session the students were placed into different levels of classes depending on their score on a placement test designed and administered by the Study Abroad 2 program at the beginning of the semester. The placement test was a 180-minute written test consisting of sections on grammar (100 questions), vocabulary (20 questions), error correction (10 sentences), listening comprehension (15 questions), reading comprehension (3 open answer questions), and writing (a 350 word composition). Each student then completed a 10-minute oral interview in Spanish with a staff member in order for the staff to verify that the student's written results corresponded accurately to their oral abilities for more accurate placement. During the intensive session Study Abroad 2 offered six levels of grammar classes and a history class that was reserved for the most advanced group of students enrolled in the intensive session (Advanced Liberal Arts). At Time 1 (early September, 2014) there were 45 participants that were recruited by the researcher from all seven levels of the intensive session and were enrolled in Liberal Arts (N=33), Communication, New Media, and Journalism (N=6), Advanced Liberal Arts (N=4), and the Teaching Development Program (N=2). The researcher visited each class, explained the study, requested the students' participation and passed around a signup sheet. Those who agreed to participate either did so immediately following the end of class or at another scheduled time during the week.

The Time 2 data were collected in mid-November, 11 weeks following the Time 1 data collection. At Time 2, 34 of the 45 participants returned to complete the study, thus immediately eliminating 11 participants from the data analysis. An additional nine participants were excluded from the study for various reasons. Three participants reported having at least one Spanish-speaking parent. However, only one reported speaking Spanish as a native language herself and two of the three

reported Spanish being spoken in the home. One of these also reported an immersion experience of five weeks of language study in the Dominican Republic. Thus, all three were eliminated from the analysis. One participant also reported that other languages, namely Hindi and Punjabi, were spoken consistently at home by her parents and was eliminated from the analysis. Two participants were eliminated due to having reported a history of trouble with hearing, one of which, it was discovered, had also participated in a pilot version of the current study several months prior to studying abroad in Seville. Three participants, who were not eliminated for any other reason, had visited a Spanish speaking country for more than one month or multiple times for a total of more than one month and were thus eliminated. In all, 20 participants from the Study Abroad 2 group were eliminated from the analysis, leaving 25 participants that were included in the analysis.

Of the 25 participants, 23 were female and two were male. All listed English as their L1 and did not speak any other L1s. The programs of enrollment for the participants were Liberal Arts (N=20), Language and Society (N=1), Communication, New Media and Journalism (N=1), and Advanced Liberal Arts (N=3). There were nine Spanish majors, 11 Spanish minors, and five who responded *neither*. Seven of the participants reported that they would take advanced courses following the intensive session at Study Abroad 2, while 15 said they would take intermediate courses and 3 did not yet know at the time. The mean grammar test score at Time 1 for the Study Abroad 2 group was 10.96 (SD = 3.68, Range = 5 - 19) out of a possible 20 points.

Regarding the participant's classroom language experience prior to studying abroad, most of the participants had taken three or more years of Spanish in high school (N=20) and at least one to two years (N=11) or three years (N=10) in middle school. Nine participants had also taken one to three years in elementary school (Table 15). None of the participants reported attending a Spanish immersion school at any point in time. At the university level, 14 participants had taken courses at the first (N=1) or

second year (N=13) level. Seven participants had taken one or more third year courses. Three had taken one or more fourth year courses. And finally, one reported having taken graduate level Spanish.

Table 15. Number of participants with each level of Spanish education prior to data collection (Study Abroad 2 group)

Education Level	0 years (# of participants)	< 1 year (# of partic.)	1-2 years (# of partic.)	3+ years (# of partic.)	Total #
Elementary school	12	4	4	5	25
Middle school	2	2	11	10	25
High school	1	1	3	20	25
University	0	0	14	11	25

Regarding the participants' reported prior experience in Spanish-speaking countries, 24 participants reported never having visited or lived in a Spanish speaking country for three weeks or more prior to arriving in Seville. One participant had spent four weeks in Costa Rica for service and language immersion.

When the learners were asked about their contact with native Spanish speakers during the two years prior to studying abroad in Seville, eight reported contact and 17 did not. The participants reported having contact with NSs from Mexico, the Dominican Republic, Guatemala, Spain (Madrid), Venezuela, Ecuador, and Los Angeles. Most reported weekly contact (N=6) and two reported daily contact. The relationships the participants had with the NSs included co-worker (N=3), teacher (N=4), friend (N=2), roommate (N=1), and significant other (N=1)⁵.

The participants also reported similar trends to the Study Abroad 1 group in terms of their use of Spanish outside of class in the year preceding the study abroad period. Table 16 shows that the majority of the participants did not speak Spanish with NSs more than a few times per year (N=15) or per month (N=6). They reported a similar frequency to the Study Abroad 1 group for the use of Spanish with non-native speakers. Like the Study Abroad 1 group, the Study Abroad 2 group reported very infrequent use of Spanish for passive language activities (e.g., TV, reading), but did report listening to

⁵ N = the total number of relationships represented

music in Spanish more than other passive activities. This was the same trend seen for the Study Abroad 1 group. Communication online and via text messaging was also done very infrequently in Spanish.

Finally, seven participants reported having studied other languages apart from Spanish and 18 had not. The languages reported included French (N=5), Italian (N=2) and American Sign Language (N=2). All reported having a beginner level of proficiency with the exception of one participant, who reported having an advanced level of proficiency in French and American Sign Language.

Table 16. Number of participants reporting Spanish use in each category prior to study abroad (Study Abroad 2 group)

Activity	A few times or less per year (# of participants)	1X/month (# of part.)	A few times per month (# of part.)	1X/week (# of part.)	A few or more times per week
Speaking Spanish with native speakers	15	0	6	3	1
Speaking Spanish with non-native speakers	14	0	4	3	4
Watching TV, videos, or movies in Spanish	16	7	2	0	0
Reading articles or books in Spanish	15	5	4	1	0
Listening to music in Spanish	6	6	4	3	6
Communicating online in Spanish	18	3	3	1	0
Texting in Spanish	19	4	2	0	0

At-home group. The participants in the At-home group were recruited by the researcher or a research assistant from multiple Spanish classes by visiting the class at the beginning or end of the period, explaining the study and then passing a signup sheet around the class. At the time of recruitment and data collection the participants were enrolled in Spanish 250 (Intermediate Spanish II, N=5), Spanish 280 (Spanish Grammar in Context, N=29), Spanish 317 (Spanish Conversation and Diction, N=2), Spanish 324 (Introduction to the Study of Hispanic Cultures, N=2), or Spanish 328 (Introduction to Hispanic Literatures, N=1). One participant from the Spanish 317 course was also enrolled in Spanish 425

(Spanish Phonetics) and Spanish 427 (The Structure of Spanish). However, no other participants reported being enrolled in any other Spanish course.

Time 2 data was collected in mid-November of 2014, 10 weeks after Time 1 data was collected. At Time 2, 34 of the original 39 participants returned to complete the study. In total, the data from 14 of the original 39 participants in the At-home group had to be excluded from the study for various reasons, including the five who did not return at Time 2. Two participants reported both Spanish and English as native languages on the background questionnaire. Five reported having at least one Spanish-speaking parent, but were also excluded for other reasons such as prior study abroad experience. Two participants reported speaking other L1s, namely Yoruba and Vietnamese. Six participants had visited a Spanish speaking country for more than one month, but five were also eliminated for other reasons and the sixth was not eliminated because his experience in a Spanish-speaking country (Peru) was only three weeks in duration and was not related to learning Spanish. Given that only five of the participants were enrolled in upper level (300/400) courses and most of them were either heritage speakers, did not return at Time 2, or had participated in an study abroad program in a Spanish-speaking country for more than one month, all five upper-level learners were eliminated. Two participants were excluded due to an error in the experiment file for one of the tasks that rendered the data uninterpretable. And finally, one participant reported having a history of hearing trouble and was excluded from the study. After all exclusions, the total number of participants remaining was 25.

Of the 25 remaining participants seven were male and 18 were female. 22 students were enrolled in Spanish 280 (Spanish Grammar in Context) and three were enrolled in Spanish 250 (Intermediate Spanish II) during the semester of data collection. 21 of the participants were planning to study Spanish as a minor and two as a major. The mean grammar test score at Time 1 for the At-home group was 10.44 (SD = 2.52, Range = 5 - 16).

Prior to attending the university, most of the participants had taken Spanish in middle (N=19) and high school (N=25), but only seven had taken Spanish since elementary school (Table 17). In high school, almost all of the participants (N=24) reported having taken three or more years of Spanish. All but one participant had taken three or more years of Spanish in high school. None of the participants had attended a Spanish immersion school at any point in their academic career. At the university level, none of the participants in the At-home group had taken any class higher than Spanish 250 or 280 prior to the semester in which data was collected.

Table 17. Number of participants with each level of Spanish education prior to data collection (At-home group)

Education Level	0 years (# of participants)	< 1 year (# of partic.)	1-2 years (# of partic.)	3+ years (# of partic.)	Total #
Elementary school	18	3	3	1	25
Middle school	6	2	11	6	25
High school	0	0	1	24	25
University	0	0	25	0	25

Of the 25 participants, only one had visited a Spanish speaking country for three weeks. However, the purpose of the trip was not to learn Spanish, but rather to take a parasitology course in various cities in Peru for three weeks.

The participants were also asked about their contact with native Spanish speakers during the two years preceding data collection. Seven participants had some degree of contact with NSs of Spanish. However, only three reported daily (N=2) or weekly (N=1) contact. All others (N=4) reported monthly contact with friends, acquaintances, teachers, co-workers, or relatives (i.e., grandparents). One participant did report having a friend from Seville, Spain but only reported annual contact and was therefore not eliminated from the study. The other places of origin of the NSs with whom the learners had contact included Mexico, Chile, Venezuela, and East Chicago.

For the part of the questionnaire regarding the amount of Spanish used outside of class during the year prior to data collection, it was determined that there was an error in the presentation of the response options in comparison to the questionnaires for the Study Abroad 1 and Study Abroad 2 groups and therefore some of the options had to be combined for the At-home group. The options *once per month* and *a few times per month* were accidentally included twice each with slightly different wording. Thus, *once a month* and *once per month* were combined and *a few times per month* and *a few times or more per month* were combined. From the data in Table 18, it is clear that the At-home group reported speaking very little Spanish with native Spanish speakers during the year preceding data collection, as did the study abroad groups. A higher number of participants in the At-home group (N=14) spoke Spanish with non-native speakers at least once per month, which is expected since they were enrolled in Spanish courses. They also reported infrequent participation in passive language-related activities (e.g., watching TV in Spanish). As for the two study abroad groups, listening to music in Spanish was the most frequent of the passive language activities. This group also communicated infrequently online and via text messaging in Spanish.

Finally, the participants were also asked about their knowledge of languages other than Spanish and English. Four participants had studied other languages, including Italian (N=1, beginner), French (N=2), German (N=1), Russian (N=1) and Arabic (N=1). Two participants reported having studied two languages other than Spanish (Italian/French, French/Arabic). All four participants reported a beginner-level knowledge of the language(s) that they had studied, with the exception of one participant who reported an advanced knowledge of French.

Table 18. Number of participants with each level of Spanish education prior to data collection (At-home group)

Activity	A few times or less per year (# of participants)	Once per month (# of participants)	A few times per month or more (# of participants)	Total #
Speaking Spanish with native speakers	20	4	1	25
Speaking Spanish with non-native speakers	11	2	12	25
Watching TV, videos, or movies in Spanish	15	7	3	25
Reading articles or books in Spanish	12	8	5	25
Listening to music in Spanish	8	10	7	25
Communicating online in Spanish	20	4	1	25
Texting in Spanish	19	1	5	25

Summary

Table 19 presents a summary of the participant data. The L2 groups are similar in terms of the number of participants (i.e., 23 or 25) and the distribution of males and females, though the At-home group has a larger number of male participants than the two study abroad groups. The large majority of participants were female in all of the L2 groups. All three L2 groups are similar in terms of the number of participants with a previous visit to a Spanish-speaking country for one month or less. Finally, the two study abroad groups reported more daily and weekly contact with NSs prior to data collection than the At-home group. In terms of age, the SEVILLE group ($M = 28.12$, $SD = 6.41$) and the Non-Aspirating group ($M = 35.4$, $SD = 3.95$) were significantly different from one another according to an independent samples t -test ($t(41) = 3.384$, $p < .01$)⁶.

Next, it is important to compare the groups' performance on the grammar test at Time 1 using independent samples t -tests. The results for the SEVILLE and Non-Aspirating groups show that the

⁶ The L2 learners were not asked about their age on the background questionnaire.

grammar test did work as it should, as both groups recorded an average of 19.24 and 19.4 out of 20 respectively, mean scores that were not significantly different from one another ($t(385) = -1.38$, $p = .168$). However, there was one test item that showed variation among the native speakers. In the sentence, ‘Yo creo que si me _____ tocado a mí dar la fiesta, no ____ dado ni la mitad de lo que ____ allí’ (I believe that if it had been up to me to throw the party, I would not have given even half of what there was), the first three blanks require the selection of a form of the auxiliary verb *haber* (have). The first two require *haber* as part of a perfect tense form. The prescriptive answers for the first two blanks are *hubiera tocado* (pluperfect subjunctive) and *habría dado* (conditional perfect) respectively. While all of the NSs selected *hubiera* for the first blank, 11 responded with *hubiera* rather than *habría* for the second blank, indicating a lack of consensus. In order to compare the other groups’ grammar test scores, independent samples *t*-tests showed that the Study Abroad 1 group’s mean grammar score was significantly higher than that of the Study Abroad 2 group ($t(670) = 7.31$, $p < .001$) and the At-home group ($t(430) = 8.61$, $p < .001$), but that scores of the Study Abroad 2 and At-home groups were not significantly different ($t(448) = 1.75$, $p = .082$). The SEVILLE group’s score was significantly higher than that of the Study Abroad 1 group ($t(226.59) = 23.54$, $p < .001$), the Study Abroad 2 group ($t(248.40) = 32.79$, $p < .001$), and the At-home group ($t(276.27) = 49.58$, $p < .001$). Likewise, the Non-Aspirating group’s score was significantly higher than all three L2 groups ($p < .001$).

Table 19. Summary of the participant groups (SA = study abroad, AH = At-home)

GROUP	TIME 1 #	TIME 2 #	SEX	GRAM. MEAN	GRAM. SD	GRAM. RANGE	# OF	# WITH
							PARTICIPANTS W/PRIOR SA EXPERIENCE	DAILY/WEEKLY NS CONTACT
SA1	32	23	3M/20F	13.04	3.69	5-19	2	7
SA2	45	25	2M/22F	10.96	3.68	5-19	1	8
AH	39	25	7M/18F	10.44	2.52	5-16	1	3
SEVILLE	33	n/a	21M/12F	19.24	0.99	16-20	n/a	n/a
NON-ASP	10	n/a	6M/4F	19.40	0.80	18-20	n/a	n/a

Instruments

General Procedure

The primary instruments for the current study included two speech perception tasks, a forced-choice identification task and a lexical decision task. These two tasks were administered to the participants in a counterbalanced manner. To accompany the lexical decision task and immediately following it, a word familiarity questionnaire was administered to control for the L2 participants' word knowledge. After the word familiarity questionnaire, participants completed a 20-item grammar test. Finally, the L2 participants completed a pre-semester background questionnaire and a post-semester native speaker contact and language use questionnaire. The two study abroad groups and the At-home group completed separate questionnaires tailored to their particular learning contexts while maintaining any questions that could be equally answered by participants in either group, allowing for a direct comparison. The native speaker groups did not complete the word familiarity questionnaire, but did complete the grammar test and a background questionnaire in that order. All instructions for the two speech perception tasks were given in Spanish both orally by the researcher or research assistant and in written form on the computer screen. This was to ensure that all participants thoroughly understood the perception tasks given that they are not like any task that the participants typically do. However, the instructions for the word familiarity questionnaire, grammar test, and background questionnaire were presented on the computer screen in Spanish (grammar test) or English (word familiarity and background questionnaires) and participants were told to read the instructions since these tasks were more easily understood by the participants.

Data was collected twice for L2 participants and once for NS participants. The first data collection occurred during at the beginning of the fall semester of 2014 (Time 1) and the second data collection occurred 10 (At-home group) or 11 weeks later (study abroad groups) in November of 2014 (Time 2). The author collected all data in Seville as well as the stateside native speaker data, while a

research assistant collected the At-home group's data at Indiana University Bloomington while the author was in Spain during Time 1. The author also collected data during T2 for the At-home group the week prior to returning to Spain in mid-November.

The Study Abroad 1 group completed all tasks in a computer lab containing eight computers that belonged to the Study Abroad 1 program using one of three types of high-quality closed headsets: Audio Technica ATH-M40FS (four sets), Bose AE2i (two sets), or Sennheiser HD280 Pro (one set). The Study Abroad 1 computer lab was a semi-quiet space, however the researcher was required by the staff to permit the students who were not participating at any given time to pass through the testing area to use the only available restroom in the building or to reach another room that was frequently needed by the students and staff. Thus, there were times when the door was opened and closed during testing, or a group of students were talking in the adjacent room. In these cases, the researcher closed the door to minimize noise in the testing room.

The Study Abroad 2 group completed all tasks in a computer lab that belonged to the Study Abroad 2 program and contained nine Windows 7 computers, using one of the same three types of headsets as the Study Abroad 1 participants. This, too, was a semi-quiet space given that it was located directly next to the main lobby of the building where students congregated between classes to talk. The glass doors to the lab were closed to diminish noise, however it was not sufficient to create a completely silent environment. Additionally, faculty, staff, and students needed to enter or pass through the computer lab to the adjacent room on occasion to use available computers for course-related purposes during the testing and the researcher was obligated by the staff to accommodate these requests. In these cases, the individuals wishing to enter were informed of the testing that was in progress, asked to be quiet, and informed that any necessary printing needed to wait until the participants had completed the task that was in progress at the time, if it was a speech perception test.

The At-home and Non-Aspirating groups completed all tasks in a computer lab or linguistics lab at Indiana University Bloomington. The headsets that the At-home group used were designed for use in language classes and specifically for listening to speech (Sanako SLH07 professional headset). The Non-Aspirating group used the Sennheiser HD-280 PRO and Audio Technica ATH-M40FS headsets. During testing, individuals who were not participating in the study were allowed to sit in the back two rows of the computer lab but were not permitted to make noise or print. During the majority of testing there was no one in the lab apart from the participants. Allowing non-participating students to be present in the lab during testing was a stipulation of the directors of the lab in exchange for reserving the lab for multiple hours at a time.

Finally, the SEVILLE group completed the tasks in various locations, including participants' homes, the home of an acquaintance of the participant, a participant's place of employment, the computer lab at one of the study abroad program facilities and in one case outside at a convenient location for the participant. All participants who did not participate in the computer lab at Study Abroad 2 completed all tasks using the researcher's personal laptop computer or another computer if available. All participants in the SEVILLE group used either the Audio Technica ATH-M40FS or Sennheiser HD280 Pro headsets to complete the perception tasks.

Forced-Choice Identification Task

A forced-choice identification task was designed following Schmidt (2011), but included some modifications. The forced-choice identification task was designed to determine how word-internal coda /s/-aspiration in Andalusia was perceived and mapped onto Spanish graphemes.

Materials. The task consisted of 180 di-syllabic non-word stimuli that were created in accordance with Spanish phonotactic rules. The syllabic structure of all stimuli was either [CVC.CV] or [CV.CV] and stress was always on the first (i.e., penultimate) syllable. In examining Schmidt's (2011) task for the creation of this task, 28 real Spanish words were found that were intended to be non-words and

were thus replaced by different words in the current task design to be non-words. Of the 180 stimuli, 42 were target stimuli. Of the 42 target stimuli, 14 contained [s] in word-internal coda position and will be referred to as the CodaS condition (e.g., [baspe]). 14 stimuli contained Andalusian post-aspiration and will be referred to as the Asp condition (e.g., [bap^he]). And 14 stimuli did not contain any coda consonant (e.g., [bape])⁷ and will be referred to as the No Coda, or NC, condition. Two different versions of the task (180 items each) were created in order to balance the stimuli in such a way that participants did not hear the same stimulus carriers in the CodaS and NC conditions and the Asp condition in a given version of the task. In other words, if a participant was assigned the first task version, he or she heard the stimuli [baspe] and [bape] but the aspirated [bap^he] appeared in the other version of the task to prevent priming (see Tables 22 and 23).

Table 20. Examples of each identification task coda condition

Coda condition	Example
CodaS	[baspe]
Asp	[bap ^h e]
No Coda	[bape]

The target stimulus carriers were controlled for the vowel preceding the word-internal syllable coda (i.e., /a e i o/) and the voiceless stop following the word-internal syllable coda (i.e., /p t k/). Following Schmidt (2011), the task included more stimuli with /i/ and /e/ preceding the coda than /a/ or /o/, and did not include /u/ as a vowel preceding the target coda position. This is because both Schmidt (2011) and the author of the current study have found in separate, informal analyses of Spanish /Vs.C/ words that /i/ and /e/ more frequently occur preceding the coda in real Spanish words than /a/ or /o/.

In both versions of the forced-choice identification task, the balance of preceding vowels and following voiceless stops was kept as consistent as possible, as shown in Table 21. Furthermore, the use of the voiceless stops /p/, /t/, and /k/ following coda /s/ was due both to Spanish frequency data

⁷ The *no coda* condition of the target stimuli was not included in Schmidt (2011), but only as a control condition with separate stimulus carriers.

presented by File-Muriel (2007), who found that Spanish coda /s/ is most frequently followed /p/, /t/ and /k/, and to the fact that Andalusian /s/-aspiration occurs when the voiceless stops follow coda /s/ (Torreira, 2006; O’Neill, 2011). The three voiceless stops were represented almost equally in both versions of the task, as shown in Table 21.

Table 21. Balance of preceding vowels and following consonants across versions of the forced-choice identification task

Version 1	Preceding Vowel	#	Version 2	Preceding Vowel	#
	/i/	5		/i/	4
	/a/	2		/a/	3
	/e/	5		/e/	4
	/o/	2		/o/	2
	Following consonant			Following consonant	
	/p/	5		/p/	4
	/t/	4		/t/	5
	/k/	5		/k/	5

Table 22. Target stimuli for version one of the forced-choice identification task

Version 1 target stimuli				
Orthographic representation	Retained [s]	No Coda	Orthographic representation	Aspirated
baspe	[baspe]	[bape]	leste	[let ^h e]
gasco	[gasko]	[gako]	goste	[got ^h e]
lespo	[lespo]	[lepo]	dasca	[dak ^h a]
lisco	[lisko]	[liko]	jasta	[xat ^h a]
mesque	[meske]	[meke]	fespe	[fep ^h e]
questo	[kesto]	[keto]	fesca	[fek ^h a]
sosca	[soska]	[soka]	despa	[dep ^h a]
bispa	[bispa]	[bipa]	fiste	[fit ^h e]
dispe	[dispe]	[dipe]	gaspo	[gap ^h o]
nista	[nista]	[nita]	mispo	[mip ^h o]
plisto	[plisto]	[plito]	nisque	[nik ^h e]
foste	[foste]	[fote]	quisque	[kik ^h e]
tespo	[tespo]	[tepo]	chesto	[tʃet ^h o]
fesco	[fesko]	[feko]	gosco	[gok ^h o]

Table 23. Target stimuli for version two of the forced-choice identification task

Version 2 target stimuli				
Orthographic representation	Retained [s]	No Coda	Orthographic representation	Aspirated
leste	[leste]	[lete]	baspe	[bap ^h e]
goste	[goste]	[gote]	gasco	[gak ^h o]
dasca	[daska]	[daka]	lespo	[lep ^h o]
jasta	[xasta]	[xata]	lisco	[lik ^h o]
fespe	[fespe]	[fepe]	mesque	[mek ^h e]
fesca	[feska]	[feka]	questo	[ket ^h o]
despa	[despa]	[depa]	sosca	[sok ^h a]
fiste	[fiste]	[fite]	bispa	[bip ^h a]
gaspo	[gaspo]	[gapo]	dispe	[dip ^h e]
mispo	[mispo]	[mipo]	nista	[nit ^h a]
nisque	[niske]	[nike]	plisto	[plit ^h o]
quisque	[kiske]	[kike]	foste	[fot ^h e]
chesto	[tjesto]	[tjeto]	tespo	[tep ^h o]
gosco	[gosko]	[goko]	fesco	[fek ^h o]

Apart from the 42 target stimuli, each of the two versions of the forced-choice identification task included 50 control stimuli of which 40 were of [CVC.CV] structure and 10 of [CV.CV] structure. The 50 control stimuli (Table 24) were unique non-word stimulus carriers (i.e., different from the target stimulus carriers) that included different consonants in word-internal coda position to determine how the participants mapped coda consonants other than the target coda consonants to their orthographic representations. The phones that occurred in coda position in the control stimuli were /l/, /r/, nasals (i.e., /n/, /m/), /f/, and No Coda (NC) (N=10 each). The justification for these comes from Schmidt (2011), who found that in two pilot studies these five codas plus <s> were the most frequently selected when [h] was heard in coda position. All 50 control stimuli occurred in both versions of the forced-choice identification task.

A total of 88 distracter stimuli targeting different consonantal contrasts and in different positions in the word were also included in the forced-choice identification task and were all of [CV.CV] structure. These were designed to distract participants from the target phone being the aspirated variant of /s/. 18 of the distracter stimuli contrasted different consonants in word-initial position, namely [p], [n], [l], [r], [f], [m], [b], [k], [x], [d], and [g] (e.g., minco, ninco, linco, pinco). Another 18 stimuli presented a contrast in voicing between [p] and [b] in word-initial position of the same or very similar non-words (e.g., pole vs. bole). Another 18 stimuli contrasted the five Spanish vowels and some of the Spanish diphthongs in word-internal position (e.g., pafa, pufa, pefa; poita, puita, piota, piuta). Another 16 stimuli contrasted voicing in word-internal [b]-[p] and [g]-[k] (e.g., mepa vs. meba, ploca vs. ploga). Finally, given that Schmidt (2011) included the Argentine dialectal contrast between word-initial and internal [tʃ] and [tʃ̺], the forced-choice identification task included a dialectal contrast relevant to Andalusian Spanish. In Andalusian Spanish deaffrication can occur, changing the affricate [tʃ] (orthographically <ch>) to the voiceless fricative [tʃ̺] in syllable-initial position in both word-initial and word-internal contexts (e.g., *coche* [ko.ʃe]). Each version of the forced-choice identification task contained five stimuli with the affricate [tʃ] in word-initial position (e.g., [tʃu.ma]), four with the affricate in word-internal syllable-initial position (e.g., [na.tʃi]), five with the fricative [tʃ̺] in word-initial position (e.g., [ʃel.to]), and four with the fricative [tʃ̺] in word-internal syllable-initial position (e.g., [ne.ʃe]). The two versions of the task were controlled so that the stimulus carriers that were presented in version one with the affricate [tʃ] were presented in version two with the fricative [tʃ̺] and vice versa, so that each participant did not hear the same stimulus carrier twice. With the exception of the 18 affricate-fricative distracters, the other 70 distracter stimuli appeared equally in both versions of the experiment (see Table 25 for all distracters). The voices of the three native talkers were counterbalanced as much as possible, resulting in an almost equal distribution of each talker's voice in each version of each experiment.

Table 24. The 50 control stimuli for the forced-choice identification task

Control coda stimuli				
/l/	/r/	nasal	/f/	NC
dalpa	derpo	lempa	tafte	bipe
telpe	firpa	lando	lofta	dape
nilto	dorte	sante	softe	fete
milte	narto	pente	defca	nite
galca	lerca	binco	tifpe	goca
belco	borco	dinca	gafca	daco
dalpo	chorde	tande	lefco	tede
lolde	garda	fendo	nofpo	side
pelbo	lergo	tinga	fafpo	goto
dolga	tirbe	gombe	pefpa	leca

Table 25. The 88 distracter stimuli (with affricate/fricative stimuli by version) for the forced-choice identification task

Distracter stimuli							
Initial C	Initial /b/-/p/	Internal V/Diphthong	Internal /b/-/p/ and /g/-/k/	[tʃ] V1	[ʃ] V1	[tʃ] V2	[ʃ] V2
pergo	pane	pafa	lepa	[tʃ]uma	[ʃ]elto	[tʃ]elto	[ʃ]uma
nergo	pame	pefa	mepa	[tʃ]ede	[ʃ]ade	[tʃ]ade	[ʃ]ede
lergo	pafo	pifa	fepa	[tʃ]ide	[ʃ]iro	[tʃ]iro	[ʃ]ide
rergo	pole	pofa	nepo	[tʃ]onte	[ʃ]omo	[tʃ]omo	[ʃ]onte
fergo	pona	pufa	lebi	[tʃ]empo	[ʃ]efi	[tʃ]efi	[ʃ]empo
minco	bane	pabi	meba	nu[tʃ]o	gui[ʃ]e	gui[tʃ]e	nu[ʃ]o
ninco	bame	peba	feba	pu[tʃ]e	ne[ʃ]e	ne[tʃ]e	pu[ʃ]e
linco	bafo	piba	tebo	na[tʃ]i	li[ʃ]a	li[tʃ]a	na[ʃ]i
rinco	bole	puba	nebo	do[tʃ]o	fa[ʃ]i	fa[tʃ]i	do[ʃ]o
pinco	boma	poba	ploca				
sergo	bango	poita	ploga				
quergo	pango	piota	fega				
mergo	bimo	puita	buca				
jergo	pimo	piuta	buga				
jinco	blina	paita	suca				
dinco	plina	piata	suga				
guinco	bonca	peita					
quinco	ponca	pieta					

Procedure. The forced-choice identification task was presented to participants via Praat (Boersma & Weenink, 2014). The experiment began with 10 practice trials. Then, participants heard the 180 randomized non-word stimuli one at a time via high-quality closed headphones and were asked to identify the orthographic representation of each word they heard from six options presented visually on the screen by using the mouse to click the box of the word they wished to select (see Appendix B). The response options for the target and control items were six non-words presented on the screen that only varied in the word-internal coda consonant (<l>, <s>, <n>, <f>, <r>, NC). The response options for the distracter stimuli consisted of phones that were similar to the target phone of a given stimulus as much as possible. For example, the response options for the stimulus [bole], for which the initial [b] was the target phone, included response options such as [bole], [pole], [gole], [dole], [nole], and [fole].

Each aural stimulus was able to be repeated only once before the repeat button disappeared. Upon selecting a response, an OK button appeared and the participant was instructed to click it to continue to the next stimulus, which then played automatically. The task included one break after the first 90 items and participants were instructed to click in order to continue. It is important to note that there was no response option labeled 'none' in the current study as there was in Schmidt (2011). There are multiple reasons for this. For one, Schmidt found very little use overall of the 'none' option among her participants. More importantly, however, is that when a participant selects 'none' it is impossible to know what they are thinking about the stimulus they have just heard. One could argue, though, that forcing a response is also problematic because a forced response when a participant does not truly want to select an available response produces misleading results. While this is true, the absence of a 'none' response informs the participant that the answer must be one of the options presented. If a participant knows a response is necessary, it is more likely that he or she will choose the option that seems best rather than simply resort to 'none' as a default for a bit of uncertainty. In total, the duration of the

forced-choice identification task was between 10 and 15 minutes, depending on the speed with which the participant completed it.

Lexical Decision Task

The lexical decision task was designed so that the listener, upon hearing an aural stimulus, was required to determine whether the stimulus was a real word or non-word in Spanish. Each real word in the task had a corresponding non-word that differed phonetically by changing a vowel or consonant to render it a non-word. Both the accuracy of responses and the response times were measured and analyzed to determine the speed and accuracy with which participants accessed lexical items in the mental lexicon under different experimental conditions.

Materials. The lexical decision task consisted of two versions, each containing 196 stimuli (88 experimental stimuli and 108 distracter stimuli). There were two experimental conditions with 44 stimuli each for a total of 88 experimental stimuli. The first experimental condition, henceforth referred to as CodaS, consisted of 22 real word and 22 non-word stimuli with a sibilant [s] in word-internal coda position preceding a voiceless stop at the onset of the following syllable. The second experimental condition, Asp, consisted of 22 real word and 22 non-word stimuli with Western Andalusian /s/-aspiration in accordance with the phonetic description of this phenomenon as presented in Chapter 2. Two versions were created so that participants would not hear the same word in both the CodaS and Asp coda conditions (e.g., [gusto] & [gut^ho]) in the same version of the task to avoid priming, which could result in faster response times (RTs) for the second appearance of a word and skew the results⁸. There were also two blocks in each version of the task so that a real word and its corresponding non-word were presented in separate blocks. It is also important to note that none of the target real word

⁸ For example, if the real word ‘gusto’ appeared in the first version in the CodaS condition, it appeared in the Asp condition [gut^ho] in the second version. However, the non-word correspondent to ‘gusto,’ which is ‘gosto,’ also appeared in the first version, but was in a separate block from ‘gusto.’

stimuli had a corresponding real word in the No Coda condition (e.g., *pasta* ‘paste’ & *pata* ‘paw’) for an important reason. Schmidt (2011) found that many participants, upon hearing an aspirated stimulus in the identification task, responded by choosing the No Coda response option. For the lexical decision task in the current study, if a CodaS also had a corresponding real word without a coda consonant (e.g., *pasta* ‘paste’ & *pata* ‘paw’), it would be impossible to know whether the listener perceived *pasta* or *pata* when hearing the aspirated variant, since both are real words. Many responses would be rendered uninterpretable. By using only real Spanish words that do not have a *no coda* counterpart (e.g., *gusto* & *gute*) this problem was avoided.

Table 26. Example of the experimental conditions of the lexical decision task

Word – CodaS	Non-word – CodaS	Word - Asp	Non-word – Asp
destino [destino]	destimo [destimo]	destino [det ^h ino]	destimo [det ^h imo]
respeto [respeto]	nepeto [nespeto]	respeto [rep ^h eto]	nepeto [nep ^h eto]

In order to choose the real words for the lexical decision task, it was necessary to account for the lexical frequency of each word, given that lexical frequency has been shown to affect response times in the lexical decision paradigm. Response times for more frequent words tend to be faster than for less frequent words (Grainger, 1990; Schilling, Rayner, & Chumbley, 1985). Additionally, since L2 learners are the target population of the current study, the real words that appear in the task must be words that at least intermediate level learners would know under normal conditions (i.e., when there is no phonological variation). Clearly, if a learner does not know a word, they will respond that it is a non-word. When this occurs it is impossible to know whether a non-word response was motivated by a lack of word knowledge or by the /s/-aspiration presented in the stimulus. To control for this effect, a word familiarity survey was administered to determine which words were known and unknown to the L2 participants and only the words that were known by the participants were submitted to the analysis.

To select frequent Spanish words and increase the probability that the participants were familiar with the words in the task, the researcher consulted *A frequency dictionary of Spanish: Core vocabulary for learners* (Davies, 2006). This frequency dictionary lists the 5,000 most frequent words in Spanish according to the 20 million word *Corpus del Español* (Davies, 2002). With five exceptions, the target real words were listed among the top 2,000 most frequent words in Spanish. According to Davies (2006), the 4,000-5,000 most frequent words in a language make up approximately 95 percent of a written text and the top 1,000 words make up approximately 85 percent of spoken words. Primarily using words that rank in the top 2,000 most frequent increases the likelihood that intermediate to advanced learners of Spanish will be familiar them. In conjunction with lexical frequency, an intermediate-level Spanish textbook, *En Contacto: Gramática en Acción* (Gill & Wegmann, 2014), was consulted in order to crosscheck words selected from the *Corpus del Español* and to search for other useable words with coda /s/ preceding a voiceless stop. The five words that were outside of the top 2,000 most frequent Spanish words or did not appear in Davies' dictionary were justified based on their appearance in the intermediate textbook.

After the words were selected, a pilot study and subsequent item analysis were conducted and determined that there were three target words in the original task design that needed to be replaced due a high number of pilot study participants reporting that they were not familiar with them. The words *escaso* 'scarce,' *ajusta* 'he/she adjusts,' and *susto* 'fright,' as well as their non-word counterparts *iscaso*, *ojusta* and *nusto* were replaced with *artista* 'artist,' *asiste* 'he/she attends' and *espejo* 'mirror' and their non-word counterparts *ortista*, *osiste* and *espajo*. Another word, *destaca* 'it stands out' and it's non-word counterpart *destaga*, and the distractor pair *persona* 'person'/*bersona* were eliminated and not replaced with another word in order to create an even number of target words for the analysis. Thus the version of the lexical decision task that was ultimately used for data collection included 44 CodaS and Asp real words, 44 corresponding non-words, 54 distractor words and 54 distractor non-words.

The 88 chosen target stimuli were di- and tri-syllabic real words and non-words that contained /s/ in word-internal coda position preceding a voiceless stop in the onset of the following syllable. Of the 176 target stimuli in both versions combined, 93 contained /s/ in the coda of an unstressed syllable (e.g., *oscuro* 'dark' [os.'ku.ro]) and 83 contained /s/ in the coda of a stressed syllable (e.g., *busca* 'he/she searches' ['bus.ka]), and the distribution of stress conditions was almost exactly equal between versions one and two. The limitations imposed by controlling for lexical frequency and intermediate learner vocabulary knowledge made it impossible to evenly distribute all vowels preceding /s/ and all voiceless stops following /s/ throughout both versions of the task. However, the vowel preceding coda /s/, the voiceless stop following coda /s/, and the stress of the syllable containing coda /s/ in the target real words were controlled to be nearly identical between versions one and two of the lexical decision task so that no matter the version of the task a participant completed, all participants heard the same number of each vowel, the voiceless stops, and stress contexts throughout the task.

Table 27. Distribution of vowels preceding /s/, /p t k/ following /s/, and stress of the syllable containing coda /s/ for the real words in both versions of the lexical decision task⁹

Version 1		Version 2	
Vowel	#	Vowel	#
/a/	8	/a/	8
/e/	38	/e/	38
/u/	6	/u/	7
/i/	27	/i/	27
/o/	9	/o/	8
Stop	#	Stop	#
/p/	20	/p/	21
/t/	46	/t/	45
/k/	22	/k/	22
Stress of coda /s/ syllable	#	Stress of coda /s/ syllable	#
Stressed	42	Stressed	41
Unstressed	46	Unstressed	47

⁹ Note that the numbers for Vowel and Consonant correspond only to the real words in each version. Their corresponding non-words were, in some cases, derived from a change of vowel preceding /s/. The numbers for syllable stress include all target non-words.

Table 28. CodaS and Asp experimental stimuli for the lexical decision task

Task version	Block 1				Block 2			
	Words	Non-words	Words	Non-words	Words	Non-words	Words	Non-words
1	<u>CodaS</u>	<u>CodaS</u>	<u>Asp</u>	<u>Asp</u>	<u>CodaS</u>	<u>CodaS</u>	<u>Asp</u>	<u>Asp</u>
	gusto	astilo	bok ^h e	pik ^h ute	gosto	estilo	bik ^h e	dik ^h ute
	éste	ravista	amit ^h að	frik ^h o	iste	revista	amot ^h ad	frek ^h o
	busca	homesto	artit ^h a	ek ^h ripir	pusca	honesto	ortit ^h a	ek ^h riþir
	costumbre	lesponde	trit ^h e	dik ^h anso	costambre	responde	trit ^h a	dek ^h anso
	puesto	isto	oet ^h e	op ^h ira	puasto	esto	oat ^h e	ap ^h ira
	presta	jiesta	ep ^h asio	upuet ^h o	prista	fiesta	ip ^h asio	opuet ^h o
	respeto	dusto	supuet ^h o	ep ^h ajo	nespeto	justo	sumuet ^h o	ep ^h ejo
	oscuro	fispano	et ^h a	et ^h eðo	osquero	hispano	it ^h a	et ^h aðo
	esposa	ascucha	det ^h ino	ek ^h opa	espisa	escucha	det ^h imo	ek ^h apa
asiste	ispera	ap ^h ekto	set ^h ema	osiste	espera	ap ^h ikto	sit ^h ema	
discurso	tisco	dap ^h ués	ot ^h ed	piscurso	disco	dep ^h ués	kot ^h umbre	
2	bosque	piscute	gut ^h o	at ^h ilo	bisque	discute	got ^h o	et ^h ilo
	amistad	frisco	ét ^h e	raþit ^h a	amostad	fresco	it ^h e	reþit ^h a
	artista	escripir	buk ^h a	omet ^h o	ortista	escribir	puk ^h a	onet ^h o
	triste	discanso	ut ^h eð	lep ^h onde	trista	descanso	kot ^h ambre	rep ^h onde
	oeste	ospira	puet ^h o	it ^h o	oaste	aspira	puat ^h o	et ^h o
	espacio	upuesto	pret ^h a	xiet ^h a	ispacio	opuesto	prit ^h a	fiet ^h a
	supuesto	espajo	rep ^h eto	dut ^h o	sumuesto	espejo	nep ^h eto	xut ^h o
	después	estedo	ok ^h uro	fip ^h ano	daspués	estado	ok ^h ero	ip ^h ano
	esta	escopa	ep ^h osa	ak ^h ucha	ista	escapa	ep ^h isa	ek ^h ucha
	destino	sestema	asit ^h e	ip ^h era	destimo	sistema	osit ^h e	ep ^h era
aspecto	osted	dik ^h urso	tik ^h o	aspicto	usted	pik ^h urso	dik ^h o	

The 108 distracter stimuli in each version of the lexical decision task were mostly di-syllabic and in some cases tri-syllabic real words (N = 54) and their corresponding non-words (N = 54). The non-words were the same as the real words with a change of one vowel or consonant to create a non-word

(e.g., queso (real – ‘cheese’) vs. quesa (non-word) or fruta (real – “fruit”) vs. frupa (non-word)). The change from real to non-word occurred in three word positions: word-initial, word-internal, or word-final. 13 non-words were created by changing a word-initial consonant (e.g., mesa vs. nesa). 14 were created by changing a word-internal consonant (e.g., acto vs. acpo). 14 were created by changing the word-final vowel (e.g., sopa vs. sopi). Finally, 13 non-word distracters were created by changing a word-initial vowel (e.g., ojo vs. ujo). The purpose of varying the location of the change was to distract the participants from focusing solely on the word-internal position in which the target coda /s/ variants appear. Table 29 shows all 108 distracter words and non-words and what was changed to make the word a non-word.

Table 29. The 108 Distracter stimuli in the Lexical Decision Task

Word	Non-word	Contrast	Word	Non-word	Contrast	Word	Non-word	Contrast
mesa	nesa	WIC	futuro	fupuro	WINC	uno	ono	WIV
queso	quesa	WFV	todo	bodo	WIC	texto	pexto	WIC
caso	gaso	WIC	solo	solu	WFV	tipo	tibo	WINC
algo	alpo	WINC	seco	secu	WFV	baño	bañu	WFV
acto	acpo	WINC	sopa	sopi	WFV	acento	ecento	WIV
vida	vipa	WINC	exacto	axacto	WIV	copa	gopa	WIC
árbol	ácbol	WINC	abril	ebril	WIV	chico	chigo	WINC
frío	frúo	WIV	cama	came	WFV	alegre	alegri	WFV
boca	bopa	WINC	carta	garta	WIC	enero	inero	WIV
bebé	pebé	WIC	mano	manu	WFV	cuerpo	tuerpo	WIC
entrar	intrar	WIV	baile	paile	WIC	foto	fota	WFV
fruta	frupa	WINC	ayuda	oyuda	WIV	hijo	hije	WFV
entre	intre	WIV	arriba	orriba	WIV	hijo	hije	WFV
gordo	gorto	WINC	dieta	pieta	WIC			
ojo	ujo	WIV	donde	donda	WFV			
oso	ofu	WINC	enfermo	anfermo	WIC			
paso	pafo	WINC	leche	lete	WINC			
pelo	belo	WIC	loco	locu	WFV			
pero	bero	WIC	otro	atro	WIV			
punto	puntu	WFV	piso	pifo	WINC			
útil	étil	WIV	poco	pocu	WFV			

Note: WIC=word-initial consonant change, WINC=word-internal consonant change, WFV=word-final vowel change, WIV=word-initial vowel change

Procedure. The lexical decision task was presented using DMDX software (Forster & Forster, 2003). DMDX is a script-based program that allows for good control of parameters and reliable response time measurements with millisecond accuracy. The task began with ten practice trials in order to familiarize the participants with the task. The ten trials consisted of words and non-words that were not related to the real task and were designed to be easy so that the task was easily understood. The participant was instructed to listen to each word and respond by pressing the Right Alt key on the keyboard if the stimulus was a real word in Spanish and the Left Alt key if it was not a real word in Spanish. The practice trials gave the participant feedback of 'good,' 'wrong' or 'too slow' with response times. The practice trials were then followed by a short break and the participant was instructed to continue to begin the rest of the task by pressing the spacebar. Feedback was automatically turned off after the practice trials. Each stimulus was preceded by a fixation cross (+) in the center of the screen that was displayed for one frame with a duration of 250 milliseconds. Following the fixation cross, the stimulus file was played. The clock that measured the response time began at the beginning of the frame that played the stimulus file so that an accurate RT could be recorded. The timeout for each trial was set at 2200 milliseconds, causing participants to have to answer quickly and respond with their initial processing of the stimuli, thus minimizing the use of metalinguistic knowledge. The 196 trials were divided into two blocks of 98 with a break between them. The participant pressed the spacebar to continue to the second block of stimuli. Importantly, both the blocks and the trials within each block were randomized, meaning that the order of presentation of the first or second block was random, as was the order of presentation of the stimuli within a block. However, the stimuli in one block were never mixed with stimuli in the other to maintain the purposeful separation of stimuli into separate blocks as previously described. The task took approximately 14 minutes to complete.

Preparation and selection of the experimental stimuli for both tasks. The aural stimuli for the two perception tasks were recorded in a soundproof recording studio on an Apple Power Mac G5

computer using a Shure SM-7b microphone and Peak LE 5 software. All stimuli were recorded at a sampling rate of 44100 Hz and 16 bit on a mono channel. The perceived loudness of all stimuli was equated by normalizing the root mean square amplitude of all the files to -23dB. Three native speakers of Andalusian Spanish, two male and one female, who were graduate students studying Hispanic Linguistics (N=2) and French Linguistics (N=1) at Indiana University, were recorded reading words from slides presented via Microsoft PowerPoint¹⁰.

One male talker and the female talker were from Seville, Spain and the second male talker was from Cádiz, Spain. The male talker from Seville was 31 years old, was born in Seville and had lived there his entire life except for approximately one year and seven months prior to recording due to having lived in the United States to attend graduate school. He reported having an advanced level of proficiency in English, an intermediate-high level of French, and advanced level of Catalan. The female talker from Seville was 25 years old and was born in Córdoba, Spain. However, she only lived in Córdoba for 2 years and then moved to Seville and has lived there for 23 years. At the time of recording she had lived in the United States for approximately seven months to attend graduate school and reported having a near-native proficiency in English (C2 European rating), intermediate-low in French and German, and beginner-high in Italian. The male talker from Cádiz was 35 years old, had lived in the United States for 11 years, and reported having an advanced level of proficiency in English, advanced proficiency in French, intermediate proficiency in Portuguese, and beginner proficiency in Haitian Creole and German. All three talkers aspirate coda /s/ in informal speech in a way that is consistent with Western Andalusian Spanish and exhibit varying degrees of aspiration or breathy voicing before a following voiceless stop that coincides with aspiration, as reported by Torreira (2006). The talkers were asked to read the words as naturally as possible, as if they were speaking with friends from their hometown, and to utilize a

¹⁰ This format was used, as opposed to a word list on paper or the use of carrier phrases, because of the large number of words to be recorded and the tendency of talkers to change the intonation of the last word in a list. Since they did not know the end point, all words were recorded with the same rising intonation.

rising intonation to avoid the word-final devoicing¹¹. Each target word, meaning CodaS, Asp, and NC words, were read twice each in order to have multiple options from which to select the experimental stimuli. All distracter words were read once each by each talker.

After the recordings were completed, the sound files were segmented and labeled using Praat and the process of selecting the experimental stimuli began. In order to select tokens for use in the tasks, the productions of each talker for each word were compared aurally and spectrographically (when necessary) in order to find the tokens that had the least aspiration/breathy-voicing before the voiceless stop and also had VOT values that were longer than VOT values for normal voiceless stops in Spanish (e.g., Rosner et al., 2000). This is because Western Andalusian speakers have been shown to produce shorter aspiration preceding the voiceless stop than most varieties in which aspiration occurs (Torreira, 2006). Western Andalusian /s/-aspiration frequently takes on a breathy character, though strident aspiration preceding the voiceless stop also occurs (O'Neill, 2011). For this reason, the tokens were not limited only to those which had zero breathiness preceding the voiceless stop, but rather included tokens with a range of breathiness (see Appendix C for acoustic measurements of the stimuli). In order to control for the duration of aspiration preceding the voiceless stop without manipulating stimuli, the token of each word that exhibited the least aspiration preceding the stop was selected provided that the VOT of the voiceless stop was also within an acceptable range for Andalusian Spanish, given that WAS voiceless stops tend to exhibit significantly longer VOT following /s/ than other dialects of Spanish (i.e., long-lag VOT). As part of the process of selecting stimuli, all experimental stimuli were analyzed acoustically by taking four measurements as shown in Table 30. Table 30 shows the mean, standard deviation, and range for the duration of the vowel preceding /s/, the duration of the coda sibilant in

¹¹ It is recognized that by recording all stimuli with rising intonation it is not possible to claim that the stimuli are representative of the norm for Western Andalusian /s/-aspiration generally, in all contexts, but that they can only be considered representative, phonetically-speaking, of the norm for words spoken with a rising, or question intonation. The drawback to this is that it is unlikely that the participants were exposed very frequently to the target words spoken with a rising intonation. This tradeoff was accepted in favor of maintaining the clarity of the words by avoiding final devoicing. The effect on the results is hypothesized to be minimal.

CodaS stimuli or the aspiration or breathy voicing preceding the voiceless stop in Asp stimuli, the duration of the stop closure preceding the release burst of the voiceless stop following coda /s/ and the VOT of the voiceless stop in the onset of the syllable following coda /s/ for all stimuli of the forced-choice identification task. Table 33 shows the same information for the stimuli from the lexical decision task, with the exception that the lexical decision task does not include a No Coda condition. Tables 31 and 34 show the descriptive statistics for VOT of each voiceless stop following /s/ according to coda condition.

Following Torreira (2006), VOT was measured from the closest zero-crossing preceding the beginning of the burst to the last downward zero-crossing before a whole first cycle of the waveform of the following vowel could be observed. The beginning point of aspiration/breathy voicing was measured from the point at which F2 showed a clear decrease in energy and there was evidence of perturbation in the waveform. The endpoint of aspiration or breathy voicing was measured with reference to the end of turbulence in high frequencies (4000-6000 Hz) in the waveform. Subsequently, the stop closure was measured from the end of the aspiration or breathiness preceding the voiceless stop to the beginning of the release burst of the voiceless top. The duration of the vowel preceding /s/ was measured from the last downward zero-crossing before the start of periodicity in the waveform to the beginning of the aspirated or breathy period. For cases in which a nasal or liquid preceded the vowel, the point in the waveform at which the waveform clearly changed was marked as the beginning of the vowel. If a diphthong preceded /s/ rather than a vowel, the whole diphthong was considered the vowel. If two vowels in hiatus preceded /s/, only the second vowel was measured and was marked beginning at the point in which the structure of the waveform that changed between the vowels.

Table 30. Descriptive statistics of the stimuli selected for the forced-choice identification task (in milliseconds).

	Condition	Mean duration (ms)	SD	Minimum	Maximum
Preceding Vowel	No Coda	130.48	30.44	85.68	232.32
	Asp	106.24	18.18	83.40	146.01
	CodaS	121.87	20.51	93.04	163.18
Sibilance/Aspiration	No Coda	0	0	0	0
	Asp	24.39	12.14	3.904	47.60
	CodaS	92.03	12.66	66.96	133.78
Stop Closure	No Coda	118.99	12.38	94.02	142.98
	Asp	137.69	23.19	94.04	188.48
	CodaS	97.31	13.63	62.83	118.67
VOT	No Coda	21.79	10.01	9.16	46.41
	Asp	55.21	15.06	31.08	94.71
	CodaS	19.17	6.41	5.96	28.55

Table 31. VOT statistics according to following consonant and coda condition for the stimuli of the forced-choice identification task (in milliseconds)¹²

Consonant	Condition	Mean VOT (ms)	SD	Minimum	Maximum
/p/	No Coda	15.80	6.72	9.16	28.62
	Asp	54.09	18.00	38.06	94.71
	CodaS	13.37	6.33	5.96	22.89
/t/	No Coda	18.39	7.21	10.10	30.57
	Asp	45.91	9.39	31.08	60.42
	CodaS	20.76	4.50	13.97	27.77
/k/	No Coda	30.25	9.35	18.05	46.41
	Asp	64.59	11.33	55.33	92.09
	CodaS	22.95	4.22	16.16	28.55

¹² For detailed acoustic measurements of all stimuli, see Appendix B

Table 32. VOT statistics according to condition for each talker (forced-choice identification task)

Talker	Condition	Mean VOT (ms)	SD	Minimum	Maximum
1 (Male, Seville)	No Coda	20.47	11.89	9.16	46.41
	Asp	56.51	19.87	31.08	92.09
	CodaS	17.95	9.38	5.96	28.55
2 (Male, Cádiz)	No Coda	22.89	8.73	11.52	38.53
	Asp	51.91	9.22	39.62	65.78
	CodaS	20.74	3.94	14.11	27.77
3 (Female, Seville)	No Coda	21.99	10.20	10.10	38.82
	Asp	56.98	15.92	32.52	94.71
	CodaS	18.84	5.23	11.45	28.45

Table 32 displays the descriptive statistics for VOT according to Condition and Talker for the forced-choice identification task. Statistical comparisons were conducted using a Kruskal-Wallis test in order to determine whether the primary acoustic cue for /s/-aspiration, length of VOT, varied according to the factors of Talker (i.e., the three speakers of Western Andalusian Spanish), Condition (i.e., Asp, CodaS, No Coda), and Consonant (i.e., /p/, /t/, or /k/ following coda /s/). Post-hoc pairwise comparisons were conducted using the Mann-Whitney U test. Medians are reported alongside means and standard deviations due to the nature of the tests (e.g., Mann-Whitney is an ordinal test).

First, the results for Talker showed that there was no significant difference in overall VOT among the three Talkers ($\chi^2(2) = .547, p = .761$). For Talker 1, the median VOT was 25.65 milliseconds ($M = 30.68, SD = 22.21$). For Talker 2, the median VOT was 25.14 milliseconds ($M = 31.85, SD = 16.25$). And for Talker 3, the median VOT was 25.77 milliseconds ($M = 33.39, SD = 21.02$). Thus, all three talkers were producing VOT with very similar acoustic measurements.

For Consonant, there was a significant difference in VOT depending on the voiceless stop that followed /s/ ($\chi^2(2) = 9.122, p < .05$). Post-hoc comparisons using the Mann-Whitney U test showed that the VOT for /p/ ($Mdn = 18.94$ ms, $M = 27.75, SD = 22.07$) was significantly lower than that of /k/ ($Mdn = 28.5$ ms, $M = 44.66, SD = 21.52, p < .01$), but that it was not significantly different from the VOT of /t/

(*Mdn* = 22.57 ms, *M* = 30.86, *SD* = 16.02, *p* = .287). The VOT for /k/ was also significantly longer than that of /t/ (*p* < .05).

Finally, using the Kruskal-Wallis test there was a significant effect of Condition (i.e., CodaS, Asp, No Coda) on VOT ($\chi^2(2) = 53.47$, *p* < .001). A Mann-Whitney U test comparing the Asp condition to the No Coda condition was significant (*U* = 15.0, *p* < .001) and showed that VOT was significantly longer in stimuli that contained /s/-aspiration (*Mdn* = 55.89 ms, *M* = 55.21 ms, *SD* = 15.06) than in stimuli that contained No Coda (*Mdn* = 19.83 ms, *M* = 21.79, *SD* = 10.01). Comparing the Asp condition to the CodaS condition (*Mdn* = 19.94 ms, *M* = 19.17, *SD* = 6.41), VOT was also significantly longer in the Asp condition (*U* = .000, *p* < .001). Finally, VOT was not significantly different between the No Coda and CodaS conditions (*U* = 356.00, *p* = .555).

Table 33. Descriptive statistics for CodaS and Asp target stimuli of the lexical decision task (in milliseconds).

	Condition	Mean duration (ms)	SD	Minimum	Maximum
Preceding Vowel	Asp	84.38	23.74	22.66	146.42
	CodaS	100.38	24.48	62.05	180.85
Sibilance/Aspiration	Asp	17.71	14.25	0.00	56.03
	CodaS	88.83	19.08	55.80	140.70
Stop Closure	Asp	124.75	26.76	48.43	178.92
	CodaS	88.11	16.26	45.24	133.13
VOT	Asp	47.89	14.53	22.66	90.25
	CodaS	18.69	8.94	6.81	52.29

Table 34. Descriptive statistics for VOT according to each voiceless stop and each condition for the lexical decision task (in milliseconds).

Consonant	Condition	Mean VOT (ms)	SD	Minimum	Maximum
/p/	Asp	36.10	10.20	20.02	53.56
	CodaS	12.57	9.72	7.06	52.29
/t/	Asp	42.88	11.16	25.79	67.69
	CodaS	17.46	5.27	6.81	29.65
/k/	Asp	64.20	13.18	46.20	90.25
	CodaS	26.84	8.80	13.85	44.48

Table 35. VOT statistics according to condition for each talker (lexical decision task)

Talker	Condition	Mean VOT (ms)	SD	Minimum	Maximum
1 (Male, Seville)	Asp	51.34	16.68	23.12	86.46
	CodaS	16.41	6.57	7.42	32.80
2 (Male, Cádiz)	Asp	49.52	15.02	20.02	90.25
	CodaS	23.51	10.33	7.36	52.29
3 (Female, Seville)	Asp	44.17	11.71	25.52	83.64
	CodaS	15.77	6.22	6.81	35.37

For the stimuli from the lexical decision task, the same Kruskal-Wallis test was conducted in order to determine whether the primary acoustic cue for /s/-aspiration, length of VOT, varied according to the factors of Talker (i.e., the three speakers of Western Andalusian Spanish), Condition (i.e., Asp, CodaS), and Consonant (i.e., /p/, /t/, or /k/ following coda /s/). Post-hoc pairwise comparisons were conducted using the Mann-Whitney U test.

First, the results for Talker (Table 35) showed that there was no significant difference in overall VOT among the three Talkers ($X^2(2) = 1.13, p = .569$). For Talker 1, the median VOT was 26.32 milliseconds ($M = 33.28, SD = 21.54$). For Talker 2, the median VOT was 29.86 milliseconds ($M = 34.72, SD = 17.99$). And for Talker 3, the median VOT was 32.09 milliseconds ($M = 32.19, SD = 17.15$).

For Consonant, there was a significant difference in VOT depending on the voiceless stop that followed /s/ ($\chi^2(2) = 22.68, p < .001$). A post-hoc comparison of VOT for /p/ and /k/ using the Mann-Whitney U test was significant ($U = 396.00, p < .001$), showing that the VOT for /p/ ($Mdn = 22.77$ ms, $M = 25.31, SD = 15.78$) was significantly lower than that of /k/ ($Mdn = 44.48$ ms, $M = 44.66, SD = 21.52$). A comparison between /p/ and /t/ was also significant ($U = 1327.00, p < .05$), showing that the VOT for /p/ was also lower than that of /t/ ($Mdn = 26.84$ ms, $M = 30.86, SD = 16.02$). Finally, a comparison between /k/ and /t/ was significant ($U = 1388.00, p < .001$), showing that VOT for /k/ was significantly longer than that of /t/.

Finally, for Condition (i.e., CodaS vs. Asp), the VOT was significantly longer ($U = 287.00, p < .001$) in stimuli that contained /s/-aspiration ($Mdn = 46.46$ ms, $M = 47.89, SD = 14.53$) than in stimuli that contained a coda sibilant [s] ($Mdn = 17.12$ ms, $M = 18.82, SD = 8.76$).

Word familiarity questionnaire. In order to supplement the lexical decision task, a word familiarity survey (WFS) was designed to determine which real Spanish words were known and which were unknown to the participants. This is important because if a word is unknown, the response a participant gives as to whether the word is a real or non-word is unreliable. The WFS was administered using the Qualtrics Research Suite and required the participants to read a list of words that only included the target real words from the lexical decision task ($N=44$) and to mark one of three response options: 1) "I know what this word means (please write what it means in English in the space below)," 2) "I have heard/seen this word before, but I do not know its meaning," or 3) "I have never seen/head this word before and do not know its meaning." The purpose of the three options rather than a dichotomous *known-unknown* response is that it is possible that a participant might recognize a word and be able to say that it is a word in Spanish without actually knowing what it means. In this way, the survey avoids ambiguous terms such as "familiar/unfamiliar," which could mean different things to different

participants. To one participants, “familiar” might mean that they know what it means, while to another participant it might mean that they have heard it somewhere before but do not know what it means. Therefore, for the analysis of the WFS responses one and two were conflated for the analysis and considered “familiar” and response three is considered “unfamiliar.”

Grammar test. The multiple-choice grammar test contained 20 multiple-choice grammatical items embedded in a contextualized story (Geeslin & Gudmestad, 2010) with each item displaying three answer choices in a drop-down menu. The test has been modified and shortened since Geeslin and Gudmestad (2010), who used a 25-item version of the test, in order to improve statistical reliability (Linford, in preparation). The 20 discrete-point items were designed to test many different properties of Spanish grammar that are taught explicitly in the classroom, including present tense verbs, subjunctive versus indicative moods, direct and indirect object pronouns, definite and indefinite article usage, possessive pronouns, preposition usage, the of the infinitive, the past tense distinction between the preterit and imperfect, the imperfect subjunctive, pluperfect subjunctive and the conditional tense. The benefit of this type of test for the current study is that it not only allows for a more reliable division into proficiency levels than course level alone could provide, but it will also show whether grammatical proficiency is or is not correlated with results for speech perception.

Language background questionnaire (Time 1). The final task that the participants completed at Time 1 was a language background questionnaire. The background questionnaire elicited information about each participant’s background with regards to learning Spanish as a second language as well as other languages studied or spoken by the participants and experience visiting Spanish-speaking countries prior to data collection. The Study Abroad and At-home groups completed background questionnaires that were similar in content, yet different due to the inherent differences between the two learning contexts. The questionnaires were designed using Linford’s (in preparation) design as a baseline and were significantly modified to suit the particular needs of the current study. Linford’s

questionnaire design, and thus the basic design of the questionnaires administered as part of the current study, are based on the Language Contact Profile (Freed, Dewey & Segalowitz 2004) and questionnaires found in Dörnyei & Taguchi (2010), Cohen, Paige, Shively, Emert & Hoff (2005). Table 36 lists the types of information elicited from the participants in the Time 1 language background questionnaire.

It is important to note that one particular aspect of the design of the questionnaire serves an important methodological purpose. Instead of the traditional five-point Likert scale for attitude questions, a sliding scale from zero to 100 was used. In other words, the participants were asked to show their agreement or disagreement with a statement about the Spanish language and/or culture by either dragging a slider toward the left, which shows disagreement, or to the right, showing agreement. The farther right or left from center that the participant drags the slider shows stronger agreement or disagreement. In this way, more variability is possible in the data by giving a number from zero to 100, thus allowing for a stronger statistical analysis¹³.

Native Speaker Contact and Language Use Questionnaire (Time 2). At Time 2, a NS contact and language use questionnaire was administered via Qualtrics to elicit the learners' self-reported contact with native speakers and language use during the semester abroad or at Indiana University. Linford's (dissertation) design was used as a baseline and was significantly modified to suit the needs of the current study. Table 37 lists the types of information elicited on the Time 2 questionnaire.

¹³ It should be noted, though, that the responses to the attitude questions were not analyzed as part of the current study.

Table 36. Background information elicited by the Time 1 background questionnaire

<p><i>Extralinguistic information</i></p> <ul style="list-style-type: none"> • Sex • Field of study/Major(s)/Minor(s) • Name of home university
<p><i>Linguistic information</i></p> <ul style="list-style-type: none"> • Participant's native language(s) • Participant's parents' native language(s) • Language(s) spoken at home among immediate family members • Other language(s) spoken by the participant and self-reported level(s) of proficiency • Self-reported comparison of the participant's speaking, listening, reading, pronunciation, and grammar abilities with those of classmates • Hearing difficulties
<p><i>Exposure information</i></p> <ul style="list-style-type: none"> • Number of years of Spanish courses in elementary, middle, and high school • Types of Spanish courses taken at different levels at their university • Specific Spanish courses taken prior to and during data collection at Indiana University (At-home group) • Spanish immersion school experience (if applicable) • Location, duration, and purpose of travel of 3 weeks or more to Hispanic countries • Frequency of Spanish exposure/use during the year prior to data collection through conversation, reading, watching visual media, listening to music, social media, and texting • Country of origin, city of origin, frequency of contact, and relationship status with any native Spanish speaker with whom the participant had relatively frequent contact with during the 2 years prior to data collection
<p><i>Attitude information</i></p> <ul style="list-style-type: none"> • Ranking of the participants' reasons for studying abroad (study abroad group) • Ranking of the participants' language learning priorities during the semester (e.g., improve pronunciation, improve grammar, read better, understand NSs, write better, understand culture, speak with a certain accent) • Ranking of the priority of certain activities during the semester (priorities differed between At-home and study abroad groups) • A sliding scale from strongly disagree to strongly agree in response to positive and negative statements about learning Spanish and different dialects of Spanish • Reasons for which the study abroad participants chose Seville over other locations

Table 37. NS contact and language use information elicited by the Time 2 questionnaire.

<p><i>Extralinguistic information</i></p> <ul style="list-style-type: none"> • Sex • Field of study/Major(s)/Minor(s)
<p><i>NS contact information</i></p> <ul style="list-style-type: none"> • Level of Spanish courses taken during the semester (beginner, intermediate, advanced) • List of courses taken during the semester including language of instruction, time per week spent in class, professor's name and professor's city of origin • Living arrangement (study abroad group) • Details regarding the frequency of conversations of 15-20 minutes or more in Spanish and English with native Spanish speakers and/or native English speakers who lived in the same residence as the participant and those who did not live in the same residence (including the sex, age, and relationship to the interlocutor and the frequency of such conversations)
<p><i>Language use information</i></p> <ul style="list-style-type: none"> • Amount of time spent per day engaging in official academic activities related to reading, writing, listening and speaking Spanish • Frequency and types of communication in English with people in the United States during the semester (study abroad group) • The percentage of use of Spanish versus English on a sliding scale from 0 to 100 percent, including all language-related activities • The percentage of time that the participant used Spanish and not English when speaking, reading (including internet, books, magazines, newspapers), or writing (including emails, homework, social media, texting) during the semester • Self-reported time spent during the average day (16 waking hours only) engaging in the following activities: Conversing in Spanish with Spanish/English speakers, conversing in English with Spanish/English speakers, listening to lectures in Spanish/ English, watching media in Spanish/English, reading in Spanish/English, and all other activities that do not involve language (e.g., shopping, dressing, walking, napping)
<p><i>Attitude information</i></p> <ul style="list-style-type: none"> • Ranking of the participants' language learning priorities during the semester as they actually ended up being (e.g., improve pronunciation, improve grammar, read better, understand NSs, write better, understand culture, speak with a certain accent) • Ranking of the priority of certain activities during the semester as they ended up being (priorities differed between At-home and study abroad groups) • A sliding scale from strongly disagree to strongly agree in response to positive and negative statements about learning Spanish and different dialects of Spanish
<p><i>Metalinguistic information about /s/-aspiration</i></p> <ul style="list-style-type: none"> • A question asking whether participants noticed or learned anything unique about how people from Seville tend to pronounce /s/ in words such as <i>estás</i> or <i>fiesta</i> • A request for a description of what the participant learned or noticed • How the participant heard or learned about /s/-aspiration • Whether the participant had met anyone from Andalusia before coming to Seville • Whether the participant believes his or her pronunciation has become more like native speakers from Seville or whether the participant has tried to speak like them.

As for the Time 1 questionnaire, a sliding scale was used in the Time 2 questionnaire not only for the attitude questions, but also for questions that elicited a percentage. For example, there were multiple questions related to the percentage of time that the participants used Spanish and English during the semester overall and also what percentage of the time each language was used when reading or using various media. A slider, unlike free response or Likert scales, allows for a more natural response based on feel.

Another design feature of the Time 2 questionnaire departs in a very important way from most study abroad questionnaires in previous studies. In the current study, rather than asking for a non-contextualized estimate of the hours per day spent using Spanish and English when speaking to Spanish and English speakers, reading, writing, listening to lectures in Spanish and English, and watching television or other media, the questionnaire for the current study elicited the number of hours on an average day spent using Spanish and English for different language-related activities (speaking, reading, writing, etc.) within the context of an average 16-hour waking period from 7:00 a.m. to 11:00 p.m. Researchers have found that when such a question is too open-ended, the total number of hours that participants list for various language use activities can easily total more than 24 hours, which is clearly unrealistic and therefore problematic for analysis (Martinsen, 2013). Even a total of 24 hours would be unrealistic because it would be ignoring the time spent sleeping, thus not giving an accurate picture of the participants' language use during the daytime. By providing the context of a 16-hour window, the questionnaire in the current study is able to have more control over the maximum value of the responses and thus a more realistic view of a learners' language use on an average day.

According to this design, each language use activity was presented with a slider from zero to 16 in increments of 0.1 hours. Thus, the participants were able to easily click and drag the slider to the desired number of hours and, in addition, numbers were displayed to the right of each slider showing how many hours the participant had selected. The instructions told the participants to ensure that the

numbers to the right of all sliders did not add up to more than 16 hours, or at least very close to 16 hours since exact calculation was time consuming and many participants did not use a calculator. Importantly, the last item served to make up the difference for all activities on an average day that were not related to language use (e.g., shopping, napping, touring etc.). Such an item is important because it allows the participants to have an outlet for any extra time and not feel that they must fit the extra time into the other sliders, which would skew the data.

Native Speaker Background Questionnaire (NSQ). The native speaker groups also completed a background questionnaire at the end of the data collection session. The NSQ elicited information regarding the NS participants' sex, age, place of origin, places in which the participant had lived and the duration of stay, other languages spoken with self-reported proficiency levels, the participant's highest education level achieved, and the job the participant had at the time of data collection.

Data Analysis

Data coding. The responses to the forced-choice identification task and lexical decision task were recorded in Praat and DMDX respectively. The experiment files were saved for each participant as text files. The text files listed the stimuli in order of presentation, which was randomized, as well as the response that each participant chose for each stimulus. A coding template was made for each version of each task so that the responses were able to be imported and partially coded automatically, with the exception of data from other sources such as the grammar test and questionnaires.

The basic coding scheme of the forced-choice identification task included the participant code for each participant, the filename for each stimulus, the six response options that appeared for each stimulus, the correct answer, the participant's answer, whether the response matched the correct response (1 or 0), the response time, the talker that recorded a given stimulus, the category of the word (i.e., target Aspirated, target CodaS, target No Coda, control, distracter), the task version, the five

acoustic measurements for all target stimuli (duration of sibilance or aspiration preceding the voiceless stop, VOT, stop closure duration, vowel duration preceding coda /s/, and vowel duration plus the sibilance or aspiration preceding the voiceless stop), preceding vowel, target consonant ([p], [t], [k], other and distracter), coda condition (i.e., NC, aspirated, [s], /l/, /r/, nasal, /f/ or control), data collection time (T1 or T2), group (Study Abroad 1, Study Abroad 2, At-home, SEVILLE, Non-Aspirating), Time 1 and Time 2 grammar test scores, the difference between the Time 1 and Time 2 grammar test scores and the Time 2 questionnaire data (to be described in detail below).

The lexical decision task was coded for participant code, the filename for each stimulus, the real or non-word status of the stimulus, the task version, the block in which the stimulus was presented, the vowel preceding coda /s/, coda condition (real or non-word in the CodaS and Asp conditions and distracters), the stress of the syllable containing coda /s/, the error rate (1 or 0 for correct or incorrect respectively), response time, group (Study Abroad 1, Study Abroad 2, At-home, SEVILLE, Non-Aspirating), data collection time (Time 1/Time 2), the number of syllables in the stimulus, and the five acoustic measurements of the target stimuli (i.e., the same measurements as the forced-choice identification task), Time 1 and Time 2 grammar test scores, the difference between Time 1 and Time 2 grammar test scores, and the Time 2 questionnaire data.

The information for each participant that was elicited from the Time 2 questionnaire was entered into the data file manually. The first variable coded was the self-reported percent use of Spanish during the semester overall. This was then broken down into three separate variables based on the percentage of Spanish use during the semester when reading, writing and speaking. Each of these four variables was coded as a number between zero and 100, representing the reported percentage. The next two variables were the number of native Spanish speakers with whom the participants reported having conversations in Spanish on a weekly and daily basis. The numbers reported under *weekly* and *daily* were treated as separate variables and were coded as a number between one and 12, which was

the maximum possible number allowed by the question. The next set of variables from the Time 2 questionnaire included the responses to the 10 parts of the question regarding their language use during a typical 16-hour day. For this item, it is important to note that most, but not all of the participants reported a total number of hours less than or equal to 16 hours, as requested in the instructions. However, a relatively small number of participants did not follow the instructions and reported a total of more than 16 hours. To control for this discrepancy, all of the responses, which were given in terms of total hours in increments of 0.1 hours, were converted to a percentage of the sum of the total hours reported for all activities. In this way, a participant who reported spending two hours out of 16 total hours speaking Spanish with native speakers of Spanish and a participant who reported spending 3.25 hours out of total 26 hours were both coded as 0.125.

Preparation of the Data for Analysis. Prior to the analysis, the data files for each task were cleaned in order to rid the files of all items that showed one of the following issues. First, any item that caused considerable difficulty for the SEVILLE group was considered problematic and excluded from the analysis. This is because the SEVILLE group, representing the target dialect, was considered the litmus test for how the tasks should operate. To determine which items of the lexical decision task were problematic for this group, accuracy was computed for all target real words (CodaS and Asp), all target non-words (CodaS and Asp), all distracter real words and all distracter non-words independently. For each group of words the mean accuracy rate and standard deviation were calculated. Any item for which the SEVILLE group's accuracy was twice the standard deviation or more below than the mean of the accuracy for a given set of items was eliminated. Importantly, only the items that met this criterion were excluded, meaning that related words were not excluded if they met the criterion for NS accuracy. For example, the aspirated non-word *escripir* was excluded, but the aspirated real word *escribir* 'to write' was not excluded because it met the criterion of exhibiting accuracy that was higher than twice the standard deviation below the mean.

Second, the same procedure was followed for the L2 data in order to exclude any real word target items with which the L2 learners were not familiar. However, rather than calculating the L2 learners' accuracy on the task as the benchmark, the percentage of L2 learners who responded that they were familiar with a given word on the Word Familiarity task (i.e., they knew the meaning of the word or had heard it before but did not know the meaning) was calculated. The mean and standard deviation were calculated and any word for which the reported percentage of respondents who were familiar with the word was equal to or lower than twice the standard deviation below the mean was excluded. Importantly, contrary to the procedure for the SEVILLE group, both the CodaS and Asp forms of each word that did not meet the criterion for inclusion in the analysis were excluded, whether or not both forms of the word did not meet the criterion. This is because if learners reported not knowing a word, they would not know it in either the CodaS or Asp conditions. This procedure was done for the Word Familiarity results both at Time 1 and Time 2. The lists of words excluded from Time 1 and Time 2 data were very similar, only differing for two items. Specifically, the words *oeste* 'west' and *presta* 'he/she lends' were under the threshold at Time 1 but had moved above the threshold by Time 2. To maintain consistency, these two words were excluded from both the Time 1 and Time 2 data. These procedures for the SEVILLE and L2 groups resulted in the exclusion of the following items from the analysis of the lexical decision task:

Table 38. Items eliminated from the analysis of the lexical decision task

Real Target	Non-word Target	Real Distracter	Non-word Distracter
discurso (CodaS & Asp)	escribir (Asp only)	acento	puntu
discute (CodaS & Asp)	piscuro (Asp only)		bañu
presta (CodaS & Asp)	piscute (Asp only)		belo
oeste (CodaS & Asp)	tisco (Asp only)		
aspira (CodaS & Asp)			

The same procedure guided the exclusion of items from the forced-choice identification task. Though no real words were part of the design of the forced-choice identification task, eliminating the

need to analyze the L2 learners' word knowledge, the SEVILLE group's data was analyzed to determine which items showed accuracy that was twice the standard deviation or more below the mean. This was done for all categories of words in the task: CodaS target items, Asp target items, No Coda target items, Control items and Distracter items. This procedure resulted in the exclusion of the following items from the analysis of the forced-choice identification task:¹⁴

Table 39. Items eliminated from the analysis of the Identification task

Target Asp	Target No Coda	Controls	Distracters
foste [fo ^h e]	bipa nita	tirbe	chefi [ʃefi] plina ploga

Response Time Data Cleaning. Response time (RT) was originally coded on a per-item basis. The response time data for the lexical decision task were cleaned by removing RTs for all incorrect responses and removing the RTs that were above or below twice the standard deviation of the mean RT for each group. As a result, 1,409 data points removed from the analysis, or 4.9 percent of the dataset for response time. This left a total of 27,130 data points to be analyzed. The RTs were then converted to log(RT) (base 10) to create a normal distribution because the RT data were right-skewed. Log(RT) was used as the dependent variable in the statistical analysis.

Analyses of Accuracy According to Group and Condition

The first analysis was a comparison of the identification accuracy of the target (CodaS, Asp, No Coda) and control coda conditions ([f], [l], [r], nasal) for all listener groups. To accomplish this, the mean percent correct for each participant in each coda condition was calculated and each participant was labeled according to group. A linear mixed effects model was run using SPSS software to determine the main effects and interaction effects of the fixed factors Group and Coda Condition on the dependent

¹⁴ Again, only these words were excluded. CodaS and Asp word counterparts of the excluded No Coda words, for example, were not excluded because they were above the accuracy threshold.

variable, which was the mean percent accuracy. Subject was included as a random effect. Post-hoc pairwise comparisons using the Sidak correction were conducted to compare the factors of Group and Coda Condition. This procedure was conducted twice, once for the Time 1 learner data (plus NS data) and once for the Time 2 learner data (plus NS data). The alpha level for all statistical analyses was .05.

Analysis of Response Type According to Group and Condition

The final analysis of the forced-choice identification task data compared the response types for each group according to the target and control coda conditions with the purpose of identifying how the participants mapped each coda condition to orthography. Cross-tabulations were calculated using Coda Condition as one variable and the coda response chosen by the participant as the other. The cross-tabulation data will be reported in a table for each group that presents the percentage and number of the total responses within each coda condition that each group chose each possible coda response. In other words, in the Asp condition, for example, the table will report how many times and what percentage of the time a group selected <s>, <f>, <l>, etcetera. The L2 groups will have two tables, one for the Time 1 response types and another for the Time 2 response types, while the NS groups will only have one table due to having one data collection time.

Analysis of Accuracy According to Group and Condition (Lexical Decision Task)

The accuracy of the participants' lexical access in each condition was analyzed by calculating each participant's accuracy rate in each condition (CodaS real words, CodaS non words, Asp real words, Asp non words, Distracter real words, Distracter non words). Thus, there were six accuracy rates calculated for each participant at each data collection time (i.e., for the L2 learners). The percent accuracy was the dependent variable in a linear mixed effects model. The model included Subject as a random effect, as well as Group and Condition as fixed effects. Apart from the main effects, post-hoc pairwise comparisons using the Sidak correction were conducted to analyze the interaction effects of the factors Group and Condition. The model was run twice separately for Time 1 and Time 2 data.

Analysis of Lexical Decision Response Time. Using $\log(\text{RT})$ as the dependent variable, a linear mixed effects model was run that included Subject and Item as random effects and Group, Condition, and Number of Syllables as fixed effects. The number of syllables was included in order to control for the inherent duration differences between two- and three-syllable words. Post-hoc comparisons using the Sidak correction were conducted to analyze interaction effects between conditions.

Analyses of the Effect of Time on L2 Learners' performance (forced-choice identification task/lexical decision task). In order to analyze the factor Time for each of the perception tasks, Time was included as a fixed factor in a linear mixed effect model that was run for each task. The model also included Subject as a random factor and Group and Condition as fixed factors. Post-hoc comparisons using the Sidak correction were used to explore interaction effects between factors.

Analyses of Extralinguistic Factors for the Study Abroad Groups (forced-choice identification task/lexical decision task). To analyze the effects of the extralinguistic factors of *language use*, *NS contact* and *grammar test score* bivariate correlation analyses were conducted using Spearman's rank correlation coefficient. The sub-factors of each of the extralinguistic factor were tested for correlation against accuracy rates at Time 1 (grammar test score only) and Time 2, as well as the accuracy differential between Time 1 and Time 2. The accuracy differential was calculated by subtracting the Time 1 accuracy rate in the aspirated condition from the Time 2 accuracy rate. This resulted in positive values for some, zero for others, and negative values for others (i.e., when the Time 2 score was less than the Time 1 score). This was then tested for correlation against the extralinguistic factors. The factor of *study abroad program* was analyzed via the linear mixed effects models above by comparing the interaction of the factors Group*Condition*Time. Specifically, the Study Abroad 1 and Study Abroad 2 groups were compared at each data collection time according to Condition to determine if there was a significant difference between the two groups.

Summary of Chapter 3

This chapter has described the method used to test the research questions that were described at the end of Chapter 2. The chapter began with a description of the participant groups: native speakers of Spanish from Seville, Spain (N=33); native speakers of Spanish who do not aspirate /s/ (N=10), study abroad learners from two different study abroad programs (N=23 & 25), and a group of learners of Spanish that remained at their home institution during the semester (N=25). This was followed by a description of the instruments used to accomplish the research goals: forced-choice identification task, lexical decision task, word familiarity questionnaire, grammar test, language background questionnaire, NS contact and language use questionnaire, and native speaker background questionnaire. The procedure that was followed for the collection of the data was described both generally and for each task specifically. The chapter concluded with a description of the data coding, data cleaning, and the descriptive and inferential statistical analyses used to analyze the data. The following chapter presents the results of the analysis that resulted from the application of the method described above.

Chapter 4 – Results

The research questions presented at the end of Chapter 2 seek to determine how L2 learners of Spanish initially (i.e., prior to exposure) categorize Andalusian /s/-aspiration orthographically and how it affects their speed and accuracy of lexical processing (research question #1), whether learners show change over time under conditions of exposure during study abroad and in the traditional (i.e., At-home) university classroom context (research question #2), and whether there is any effect of the study abroad learners' grammatical proficiency (i.e., grammar test score), study abroad program, contact with NSs and target language use during the semester abroad on their categorization and lexical processing related to /s/-aspiration (research question #3). To answer the three research questions, this chapter describes the results of the forced-choice identification task and lexical decision task for both the native and non-native listener groups. The results of the forced-choice identification task as they answer each research question will be presented first, followed by the results of the lexical decision task.

In order to answer the first research question regarding learners' initial categorization of /s/-aspiration, the chapter begins with the presentation of the accuracy data for the forced-choice identification task, including the native speaker and L2 learner data. This section will compare accuracy of orthographic identification in each of the seven coda conditions (i.e., [s], aspirated, *No Coda*, [f], [l], [r], nasals ([m]/[n])) among the groups (SEVILLE, Non-Aspirating, Study Abroad 1, Study Abroad 2, At-home), presenting Time 1 and Time 2 data for the L2 groups separately. Then, to answer the second research question as it relates to the identification patterns over time during study abroad (Study Abroad 1 and Study Abroad 2 groups) and in the at-home context (At-home group), an analysis of each group's identification accuracy in each coda condition at Time 1 and Time 2 will be presented determine whether there is a significant effect of exposure to /s/-aspiration in the input during study abroad over time. This will also include a brief comparison of the two study abroad groups to answer part of research question three regarding whether there are differences among the learners in the two study abroad

programs. Next, the responses given (i.e., orthographic mappings) by each group for each coda condition at Time 1 and Time 2, will be presented to show how the participants identified /s/-aspiration orthographically when it was not identified as <s>. Then, to answer the third research question as it relates to the forced-choice identification task, the relationships between the accuracy results and the independent variables of grammar test score, contact with native speakers, target language use are explored using correlation analyses.

Following the results of the forced-choice identification task, the accuracy results for the lexical decision task will be presented in the same order, comparing the accuracy in each condition (i.e., [s], aspirated, distracters) among the listener groups (i.e., SEVILLE, Non-Aspirating, Study Abroad 1, Study Abroad 2, At-home) at Time 1 and Time 2 separately (research question #1). This will include a comparison of the accuracy of the five groups in each condition and then a comparison of accuracy in each condition within each group separately. Then, the L2 listener groups at Time 1 and Time 2 will be compared to analyze the effect of exposure to /s/-aspiration during study abroad versus a lack of exposure in the at-home context (research question #2). This will also include an analysis comparing the learners in the two study abroad programs (sub-question of research question #3). The section detailing accuracy on the lexical decision task will end with the results of the correlation analyses of the independent variables of grammar test score, contact with native speakers and target language use (research question #3). Finally, the results of the lexical decision task for response time (RT) will be presented, comparing the listener groups' RT according to condition and at each data collection time separately (for the L2 groups), followed by a comparison of the L2 groups across time. The RT data is presented to answer parts of research questions one and two as they relate to the speed of lexical processing both at Time 1 and then over time.

Results of the Forced-Choice Identification Task

Identification Accuracy for each Coda Condition according to Group (Time 1)

To analyze the accuracy of each listener group in each coda condition, the mean accuracy and standard deviations were calculated for each individual in each coda condition (No Coda, Asp, CodaS, [r], [l], [f], nasal) for each listener group at both data collection times (i.e., for the L2 groups). A linear mixed effects model was run to evaluate the change in accuracy by coda condition between groups, comparing the NS data to the L2 learners' Time 1 data. Subject was included as a random effect. Fixed effects were Group (SEVILLE, Non-Aspirating, Study Abroad 1, Study Abroad 2, At-home) and Coda Condition (No Coda, Asp, CodaS, [f], [l], [r], and nasals) along with their interactions. The results of the model showed significant main effects for Group ($F(4, 111) = 52.183, p < .001$) and Condition ($F(5, 666) = 522.913, p < .001$), as well as a significant interaction between Group and Condition ($F(24, 666) = 23.898, p < .001$). Post-Hoc comparisons using the Sidak correction were run to determine the significance of the comparisons between each group in each coda condition.

The mean accuracy and standard deviations for all identification responses for the native speaker groups and the Time 1 L2 learner data are presented in Table 40.¹⁵ What is immediately evident is that all of the listener groups' mean accuracy was very high (93.0% to 100%) for the *No Coda*, CodaS, [l] and nasal coda conditions. Post-hoc pairwise comparisons revealed no significant differences between any of the listener groups in these four coda conditions. There were three conditions, however, that did evidence significant differences between some of the groups: Asp, [f] and [r].

¹⁵ Note that the *control no coda* ("Control NC") condition and distracters ("Dis") that are presented in Table 40 were excluded from the analysis. The *control no coda* condition consisted of stimuli with no coda consonant that were not the same stimulus carriers as the CodaS ([s]) and Asp stimuli. This condition was excluded from the analysis because of its similarity to the target NC condition both in terms of content and results. Both the target and control NC conditions exhibited very high accuracy rates among all groups. Therefore, it was deemed unnecessary to analyze the *control no coda* items in order to increase the statistical power of the analysis by eliminating an unnecessary level of the independent variable.

Table 40. Overall Identification accuracy (%) by Group and Coda Condition (Time 1) (SA = study abroad, AH = At-home)

Group		No Coda	Asp	CodaS	[r]	[l]	nasal	[f]	Control NC	Dis
AH	Mean	93.0	0.30	99.4	90.2	96.4	99.6	79.2	94.4	92.4
	<i>SD</i>	6.80	1.54	1.98	18.24	4.89	2.00	18.01	8.21	3.27
SA1	Mean	94.7	18.1	99.1	93.7	95.2	99.6	81.3	94.8	93.7
	<i>SD</i>	8.24	17.91	3.27	9.95	8.98	2.09	20.07	7.90	5.21
SA2	Mean	96.4	16.7	100.0	93.3	96.0	99.6	75.2	94.8	94.8
	<i>SD</i>	4.88	21.73	0.00	11.57	6.46	6.32	21.82	8.23	3.27
SEVILLE	Mean	99.5	74.5	99.8	99.7	99.7	100.0	99.7	99.7	99.0
	<i>SD</i>	1.17	31.04	1.24	1.93	1.74	0.00	14.66	1.74	1.50
NON-ASP	Mean	100.0	26.1	100.0	98.9	100.0	100.0	96.0	100.0	96.2
	<i>SD</i>	0.00	26.85	0.00	3.51	0.00	0.00	5.16	0.00	4.20

First, for the Asp condition, the SEVILLE group's accuracy rate ($M = 74.5\%$, $SD = 31.04$) was significantly higher ($p < .001$ for all pairwise comparisons) than that of the Non-Aspirating group ($M = 26.1\%$, $SD = 26.85$), Study Abroad 1 group ($M = 18.1\%$, $SD = 17.91$), Study Abroad 2 group ($M = 16.7\%$, $SD = 21.73$) and the At-home group ($M = 0.3\%$, $SD = 1.54$). Likewise, the At-home group's accuracy rate in the Asp condition ($M = 0.3\%$, $SD = 1.54$) was significantly lower than that of all other groups ($p < .001$). The Study Abroad 1 and Study Abroad 2 groups were not significantly different from one another ($p = 1.00$). Finally, the Non-Aspirating group was not significantly different from the Study Abroad 1 ($p = .510$) or Study Abroad 2 ($p = .264$) groups.

It is important to note that, in the Asp condition, the SEVILLE group's accuracy rate of 74.5 percent was lower than expected for a group of native speakers from the target dialect region and this was driven by a group of eight individuals who answered incorrectly for 50 percent or more of the items in this condition. The incorrect responses of these eight individuals made up 80 of the 114 total incorrect responses by Seville native speakers in the Asp condition, or 70.1 percent of the incorrect responses. Of the eight respondents who were inaccurate in the Asp condition over 50 percent of the time, seven chose the response <f> categorically or very near categorically when they heard an aspirated stimulus. In other words, the 74.5 percent rate of accuracy in this condition was not caused by

a large number of incorrect responses distributed evenly among the participants of this group, nor by any particular stimuli that exhibited a particularly low rate of correct responses, but rather by a small subset of this group that treated the task the same way by preferring the <f> response. When the scores of these individuals are not included in the accuracy rate for the SEVILLE group, the accuracy for this group in the Asp condition rises to 92.4 percent. However, because this group did not consist of only one or two individuals and it is clearly a subset of the native speaker group that performed similarly to one another by preferring the <f> response, their results were not discarded from the analysis, but rather will be discussed in Chapter Five in order to hypothesize about the reasons behind the pattern attested.

In the [f] condition, the SEVILLE group's accuracy ($M = 99.7\%$, $SD = 14.66$) was significantly higher than all three L2 groups ($p < .001$ for all comparisons), but was not significantly different from the Non-Aspirating group ($M = 96.0\%$, $SD = 5.16$, $p = .991$). The Non-Aspirating group's accuracy in the [f] condition, too, was significantly higher than the Study Abroad 1 ($M = 81.3\%$, $SD = 20.07$, $p < .01$), Study Abroad 2 ($M = 75.2\%$, $SD = 21.82$, $p < .001$) and At-home ($M = 79.2\%$, $SD = 7.64$, $p < .01$) groups' accuracy. Thus, both NS groups' accuracy for [f] was significantly higher than all three L2 groups. The Study Abroad 1 and Study Abroad 2 means were not significantly different ($p = .515$), and the At-home group, too, was not significantly different from the Study Abroad 1 ($p = .999$) and Study Abroad 2 ($p = .921$) groups.

In the [r] condition, the only pairwise comparison that reached statistical significance was the At-home group compared to the SEVILLE group. The SEVILLE group's mean accuracy ($M = 99.7\%$, $SD = 1.74$) was significantly higher ($p < .05$) than the At-home group's accuracy ($M = 90.2\%$, $SD = 18.24$). All other pairwise comparisons, though, did not reach significance ($p = .339$ to 1.00).

Summary. The results of the linear mixed effects model for the Time 1 data show that for most of the coda conditions all listener groups were not significantly different from one another. The

exceptions were the Asp, [f] and [r] coda conditions. For the Asp condition, the SEVILLE group's accuracy was significantly higher than all other groups and the At-home group was by far the lowest, being significantly lower than all other groups. The two study abroad groups, though, were statistically equal to the Non-Aspirating group for the perception of Asp stimuli and were significantly more accurate at identifying /s/-aspiration as <s> than the At-home group. The group differences for the [f] and [r] conditions were not as drastic as for the Asp condition. Both NS groups were significantly more accurate in the [f] condition than all three L2 groups and the SEVILLE group was significantly more accurate in the [r] condition than the At-home group, but was not different from the other three groups.

Identification Accuracy for Each Listener Group according to Coda Condition (Time 1)

Given that the variables Group and Coda Condition showed a significant interaction, the accuracy of each group according to coda condition will be presented separately. First, the NS data will be presented as a baseline for comparison to the L2 groups. Following the NS data, the data for each L2 group at Time 1 will be presented. The data presented below are from the same linear mixed effects model described above, but different post-hoc comparisons.

SEVILLE group. For the SEVILLE group, the only significant difference according to the linear mixed effects model was between the Asp condition and all other coda conditions ($p < .001$ for all pairwise comparisons), while the accuracy in the other coda conditions was not significantly different among them ($p = 1.00$ for all pairwise comparisons).

Non-Aspirating group. The results for the Non-Aspirating group are the same as for the SEVILLE group in that the mean accuracy for the Asp condition ($M = 26.1\%$, $SD = 26.85$) was significantly lower than for all other coda conditions ($p < .001$ for all pairwise comparisons), while the other coda conditions were not significantly different from one another ($p = 1.00$).

Study Abroad 1 group (Time 1). The Study Abroad 1 group at Time 1 also exhibited a low accuracy rate for the Asp condition ($M = 18.1\%$, $SD = 17.91$), which was significantly lower than the accuracy rate in all other coda conditions ($p < .01$ for all pairwise comparisons). Additionally, the accuracy among the CodaS ($M = 94.7\%$, $SD = 8.24$), [l] ($M = 95.2\%$, $SD = 8.98$), [r] ($M = 93.7\%$, $SD = 9.95$) and nasal ($M = 99.6\%$, $SD = 2.09$) coda conditions was not significantly different ($p = .838$ to 1.00). However, contrary to the results for the NS groups, the Study Abroad 1 group's accuracy in the [f] condition ($M = 81.3\%$, $SD = 20.07$) was significantly lower than the *No Coda* ($p < .01$), [s] ($p < .001$), [l] ($p < .01$), [r] ($p < .01$) and nasal ($p < .001$) conditions, but was still significantly higher than the Asp condition ($p < .001$).

Study Abroad 2 group (Time 1). The Study Abroad 2 group's accuracy was significantly lower for the Asp ($M = 16.7\%$, $SD = 21.73$) and [f] conditions ($M = 75.2\%$, $SD = 21.82$) than all other conditions, a result that was significant ($p < .001$) for all pairwise comparisons. The Study Abroad 2 group also patterned very similarly to the Study Abroad 1 group for the CodaS ($M = 100.0\%$, $SD = 0.00$), [r] ($M = 93.3\%$, $SD = 11.57$), [l] ($M = 96.0\%$, $SD = 6.46$), nasal ($M = 99.6\%$, $SD = 6.32$) and *No Coda* ($M = 96.4\%$, $SD = 4.88$) conditions, for which accuracy was not significantly different ($p = .570$ to 1.00).

At-home Group (Time 1). Overall, the At-home group patterned very similarly to the two study abroad groups for all coda conditions in terms of the significance of the pairwise comparisons. The At-home group's accuracy rate in the Asp condition ($M = 0.3\%$, $SD = 1.54$) was significantly lower than all other coda conditions at $p < .001$. The accuracy in the [f] condition ($M = 79.2\%$, $SD = 18.01$), was significantly higher than the accuracy in the Asp condition ($p < .001$) and lower than the *No Coda* ($M = 93.0\%$, $SD = 6.80$, $p < .001$), [r] ($M = 90.2\%$, $SD = 18.24$, $p < .05$), CodaS ($M = 96.4\%$, $SD = 4.88$, $p < .001$), [l] ($M = 96.4\%$, $SD = 4.88$, $p < .001$), and nasal ($M = 96.4\%$, $SD = 4.88$, $p < .001$) coda conditions. The coda conditions other than the Asp and [f] conditions were not significantly different from one another ($p = .077$ to 1.00).

Summary. Figure 7 presents the overall accuracy for all groups according to coda condition. The linear mixed effects model showed significant main effects for Group, Condition, and a significant interaction effect of Group*Condition. This means that accuracy among the groups varied significantly, that accuracy among the conditions varied significantly, and that the accuracy of the groups varied significantly depending on Condition (i.e., interaction). In sum, all groups showed no significant difference of mean accuracy for the CodaS, [l], [r], nasal and *No Coda* conditions. On the other hand, all groups' accuracy in the Asp condition was significantly lower than all other coda conditions. In this condition, the SEVILLE group's accuracy was significantly higher than that of all other groups and the At-home group's accuracy was significantly lower than all other groups at Time 1, while the study abroad groups and Non-Aspirating group were statistically equal. For the [f] condition, all three L2 groups patterned similarly but their accuracy was significantly lower than that of the NS groups.

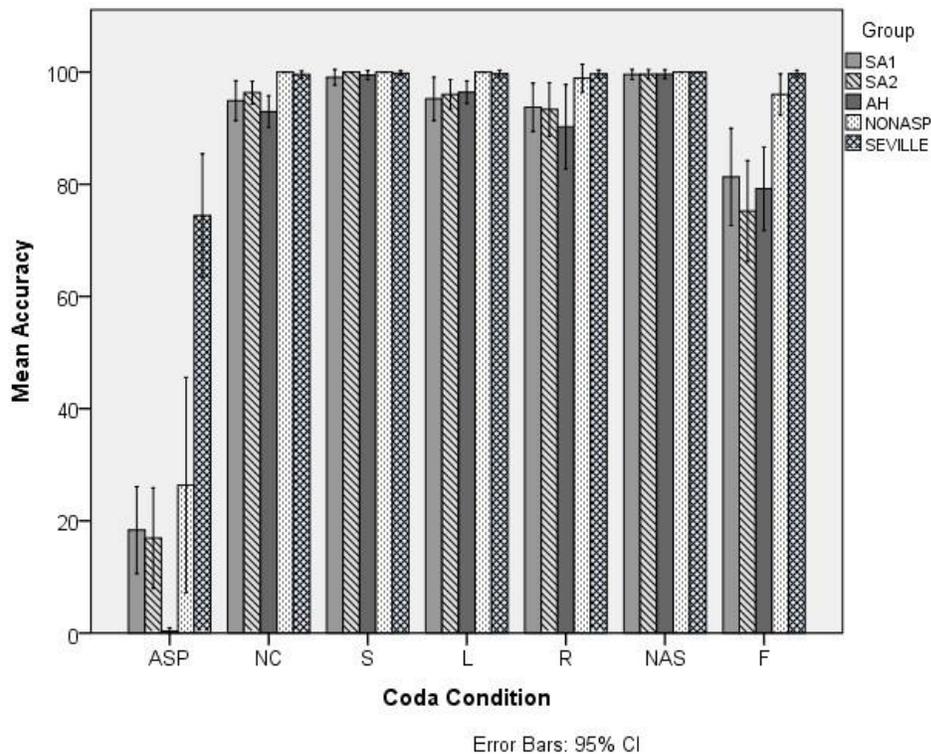


Figure 7. Summary of mean Identification accuracy (%) by Group Coda Condition at Time 1 (N=116)

Identification Accuracy for each Coda Condition according to Group (Time 2)

In order to analyze the L2 learner groups' accuracy at Time 2, the mean accuracy and standard deviations were calculated for each participant in each coda condition according to group. The NS data is repeated here as a comparison to the Time 2 results for the L2 learner groups. It is important to make clear that there was not a second data collection for the NS groups. As for the Time 1 data, a linear mixed effects model was run with the accuracy rate in each coda condition for each participant as the dependent variable, Group and Condition as the fixed factors, and Subject as a random factor. The test again showed significant main effects for Group ($F(4, 111) = 40.052, p < .001$) and Condition ($F(6, 666) = 399.100, p < .001$), as well as a significant interaction between Group and Condition ($F(24, 666) = 19.108, p < .001$). Post-Hoc comparisons using the Sidak correction were run to determine the significance of the comparisons between each group in each coda condition.

The mean accuracy and standard deviations for all Identification responses for the native speaker groups and the Time 2 learner data are presented in Table 41. Overall, the Time 2 data patterned very similarly to the Time 1 data in terms of the significance of the pairwise comparisons of the groups within each coda condition. The Time 2 data show that the mean accuracy of each group was very high (94.6% to 100.0%) for the No Coda, CodaS, [l] and nasal conditions. Post-hoc pairwise comparisons revealed no significant differences between any of the listener groups in these four coda conditions ($p = .690$ to 1.00). Notwithstanding, the Asp, [f] and [r] conditions again showed significant differences between some of the groups.

Table 41. Overall Identification accuracy (%) by Group and Coda Condition (Time 2) (SA = study abroad, AH = At-home)

Group		No Coda	Asp	CodaS	[r]	[l]	[m] [n]	[f]
AH	Mean	95.9	0.9	99.4	90.7	97.2	99.2	86.0
	<i>SD</i>	6.19	3.38	1.98	16.25	4.58	2.77	17.56
SA1	Mean	94.6	35.8	100.0	95.7	98.3	99.6	87.0
	<i>SD</i>	7.94	27.63	0.00	8.03	3.88	2.09	14.60
SA2	Mean	97.2	38.2	100.0	93.3	98.4	100.0	86.0
	<i>SD</i>	4.72	30.46	0.00	9.07	3.74	0.00	12.58
SEVILLE	Mean	99.5	74.5	99.8	99.7	99.7	100.0	99.7
	<i>SD</i>	1.73	31.04	1.24	1.93	1.74	0.00	1.74
NON-ASP	Mean	100.0	26.1	100.0	98.9	100.0	100.0	96.0
	<i>SD</i>	0.00	26.85	0.00	3.51	0.00	0.00	5.16

First, for the Asp condition, the SEVILLE group's accuracy rate ($M = 74.5\%$, $SD = 31.04$) was significantly higher ($p < .001$) than that of the Non-Aspirating group ($M = 26.1\%$, $SD = 26.85$), Study Abroad 1 group ($M = 35.8\%$, $SD = 27.63$), Study Abroad 2 group ($M = 38.2\%$, $SD = 30.46$) and the At-home group ($M = 0.9\%$, $SD = 3.38$). The At-home group's accuracy rate in the Asp condition was still significantly lower than all other groups ($p < .001$). Interestingly, the Study Abroad 1 and Study Abroad 2 groups increased accuracy in the aspirated condition at Time 2 and surpassed the Non-Aspirating group's mean. However, the Study Abroad 1 group and the Non-Aspirating group were still not significantly different ($p = .327$) and the Study Abroad 2 group's accuracy rate compared to the Non-Aspirating group was approaching significance ($p = .071$).

In the [f] condition, the SEVILLE group's accuracy ($M = 99.7\%$, $SD = 14.66$) was significantly higher than the Study Abroad 1 ($M = 87.0\%$, $SD = 14.60$, $p < .01$), Study Abroad 2 ($M = 86.0\%$, $SD = 12.58$, $p < .001$) and At-home ($M = 86.0\%$, $SD = 17.56$, $p < .001$) groups. One difference from Time 1 is that at Time 2 the Non-Aspirating group's accuracy in the [f] condition was not significantly different from that of any other group ($p = .210$ to $.992$). This is because all three L2 groups' accuracy increased in the [f] condition at Time 2 while the other control conditions remained relatively stable. Thus, only the SEVILLE group's accuracy for [f] was significantly higher than all three L2 groups. The Study Abroad 1 and Study

Abroad 2 groups were not significantly different from one another ($p = 1.00$) and the At-home group was not significantly different from Study Abroad 1 ($p = .354$) or Study Abroad 2 ($p = .210$).

In the [r] condition, as at Time 1, the only pairwise comparison that reached statistical significance was the At-home group compared to the SEVILLE group. The SEVILLE group's mean accuracy ($M = 99.7, SD = 1.74$) was significantly higher ($p < .05$) than the At-home group's accuracy ($M = 90.7, SD = 16.25$). All other pairwise comparisons, though, did not reach significance ($p = .354 - 1.00$).

Summary. The results of the linear mixed effects model for the Time 2 data show very similar patterns to the Time 1 data. For the CodaS, [l], nasal and No Coda conditions all listener groups were not significantly different from one another. The exceptions were the Asp, [f] and [r] conditions. For the Asp condition, the SEVILLE group's accuracy was still much higher than that of all other groups and the At-home group was the lowest of all the groups. The two study abroad groups, though, were not different from the Non-Aspirating group for the identification of Asp stimuli, despite being more accurate than the Non-Aspirating group at Time 2. The Study Abroad 2 group's accuracy, though, was almost significantly higher than the Non-Aspirating group for Asp stimuli. For the [f] condition, the SEVILLE group was significantly more accurate than all three L2 groups, while the Non-Aspirating group was not, given that the three L2 groups' accuracy in the [f] condition increased at Time 2. Finally, the SEVILLE group was still significantly more accurate in the [r] condition than the At-home group at Time 2, but was not different from the other three groups, which were not significantly different from one another. Next, each L2 group's results according to coda condition at Time 2 will be presented separately.

Identification Accuracy for each Listener Group according to Coda Condition (Time 2)

Study Abroad 1 group (Time 2). The Study Abroad 1 group's data at Time 2 was similar to Time 1 in terms of the significance of the pairwise comparisons between coda conditions, with some differences. The accuracy in the Asp condition ($M = 35.8\%, SD = 27.63$), though an increase is observed

(to be discussed later), was still significantly lower than the accuracy rate in all other coda conditions ($p < .01$ for all pairwise comparisons). Additionally, accuracy for the CodaS ($M = 100.0\%$, $SD = 0.00$), [l] ($M = 98.3\%$, $SD = 3.88$), [r] ($M = 95.7\%$, $SD = 8.03$) and nasal ($M = 99.6\%$, $SD = 2.09$) coda conditions were not significantly different from one another ($p = .913$ to 1.00). In the [f] condition, accuracy ($M = 87.0\%$, $SD = 14.60$) was significantly lower than in the CodaS ($p < .01$), [l] ($p < .05$), and nasal ($p < .01$) conditions, but was still significantly higher than the aspirated condition ($p < .001$). On the other hand, contrary to the Time 1 data, [f] accuracy was not significantly different from the No Coda ($p = .478$) and [r] ($p = .227$) conditions since the Study Abroad 1 group's accuracy in the [f] condition increased more at Time 2 than accuracy in the No Coda and [r] conditions did.

Study Abroad 2 group (Time 2). The results for the Study Abroad 2 group at Time 2 also generally followed Time 1 patterns, with some differences. Again, accuracy rates for CodaS ($M = 100.0\%$, $SD = 0.00$), [l] ($M = 98.4\%$, $SD = 3.74$), [r] ($M = 93.3\%$, $SD = 9.07$) and nasal ($M = 100.0\%$, $SD = 0.00$) coda conditions were not significantly different from one another ($p = .620$ to 1.00). Accuracy in the Asp condition ($M = 38.2\%$, $SD = 30.46$), though, was still significantly lower than all other coda conditions, a result that was significant at the $p < .001$ level for all pairwise comparisons. This was true even though an increase was observed compared to Time 1 (to be discussed later). For the [f] condition ($M = 86.0\%$, $SD = 12.58$), accuracy was significantly higher than accuracy in the Asp condition ($p < .001$) and lower than accuracy in the CodaS ($p < .01$), [l] ($p < .01$), nasal ($p < .01$) and No Coda ($p < .05$) conditions. The [f] and [r] conditions, though, were not significantly different ($p = .443$), which is a result that differs from Time 1. This is because the Study Abroad 2 group's accuracy in the [f] condition increased by 10.8 percent at Time 2 while the [r] condition remained the same.

At-home group (Time 2). Overall, the At-home group at Time 2 also patterned very similarly to Time 1 for the target coda conditions. The At-home group's accuracy in the Asp condition ($M = 0.9\%$, $SD = 3.38$), again, was significantly lower than for all other coda conditions at $p < .001$ and very little change

was observed (to be discussed later). The accuracy in the [f] condition ($M = 86.0\%$, $SD = 17.56$) was significantly higher than the accuracy in the Asp condition ($p < .001$) and lower than the accuracy in the CodaS ($M = 99.4\%$, $SD = 1.98$, $p < .01$), [l] ($M = 97.2\%$, $SD = 4.58$, $p < .05$), and nasal ($M = 99.2\%$, $SD = 2.77$, $p < .01$) coda conditions. The accuracy in the [f] condition, though, was not significantly lower than in the No Coda condition ($M = 95.9\%$, $SD = 6.19$) at Time 2, but was approaching significance ($p = .055$). Also contrary to Time 1 results was the fact that the [f] accuracy rate was not significantly lower than [r] ($M = 90.2\%$, $SD = 18.24$, $p = .974$) at Time 2. As for the two study abroad groups, the [f] condition showed an increase in accuracy at Time 2 while the [r] condition showed little change. Like Time 1, though, the [r], [l], [s], nasal and No Coda conditions were not significantly different from one another ($p = .164$ to 1.00).

Summary. Figure 8 presents the overall accuracy data for all groups according to coda condition, including only the Time 2 data for the L2 groups. The figure reflects the Group*Condition interaction in that all groups showed no significant differences in mean accuracy for the CodaS, [l], [r] and No Coda conditions. However, there were group differences in the Asp and [f] conditions. The accuracy of all groups in the Asp condition was significantly lower than that of all other coda conditions, but there were also group differences in that the study abroad groups were significantly more accurate than the At-home group and the SEVILLE group was significantly more accurate than all other groups. For the L2 groups, the [f] condition continued to be more difficult than the other control coda conditions, though there was an increase in accuracy in the [f] condition for all three L2 groups at Time 2. This changed the relationship between the [f] condition and other conditions in terms of significance. For the Study Abroad 1 group, accuracy for [f] was significantly lower than the CodaS, [l] and nasal conditions. For the Study Abroad 2 group it was significantly lower than the No Coda, CodaS, [l] and nasal conditions. And finally, for the At-home group, accuracy for [f] was significantly lower than for the CodaS, [l] and nasal conditions. Given the increase in accuracy for [f] for all three groups and the relatively stable position of

accuracy for [r] over time, [f] was no longer significantly lower than [r] for any of the L2 groups at Time 2.

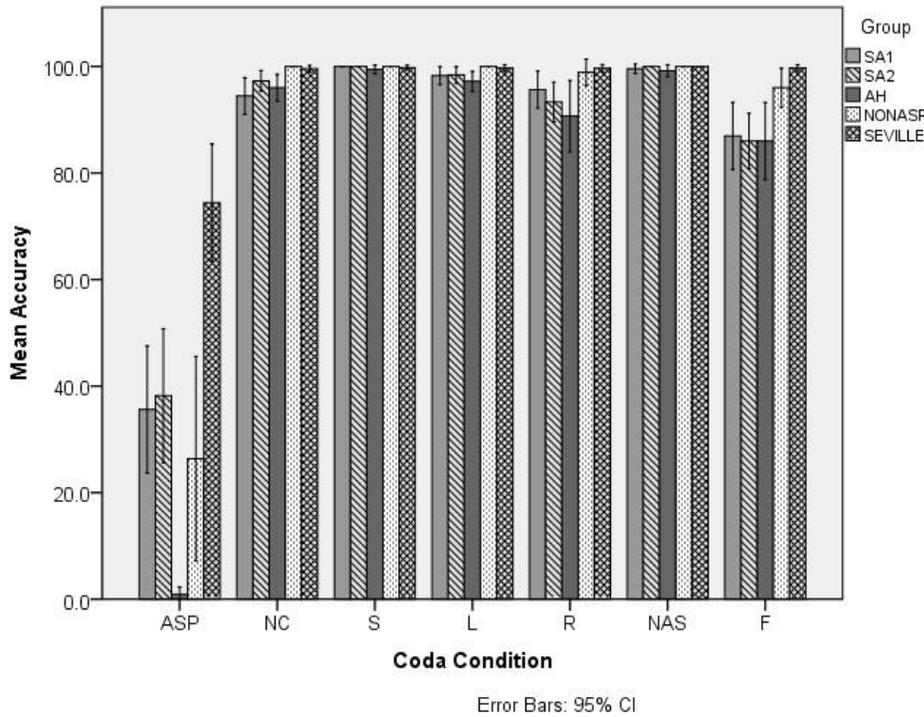


Figure 8. Summary of mean Identification accuracy by Coda Condition – All groups at Time 2 (N=116)

Analysis of the Effect of Time (i.e., Exposure) on the L2 Learners’ Identification Accuracy

Now that the data from Time 1 and Time 2 have been presented separately, revealing the overall patterns of each group in each coda condition at each data collection time (i.e., twice for the L2 groups and once for the NS groups), the effect of the variable Time on the L2 learners’ performance on the forced-choice identification task will be examined in order to test whether exposure to Andalusian /s/-aspiration over time played a significant role for the L2 groups’ accuracy in the Asp condition compared to all other coda conditions. To accomplish this, a linear mixed effects model was run to evaluate the change in accuracy by coda condition according to time and between groups. Fixed effects included were Time (Time 1, Time 2), Group (Study Abroad 1, Study Abroad 2, At-home), and Coda

Condition (Asp, CodaS, No Coda, [f], [r], [l], nasal) along with their interactions. Subject was included as random effect. The model showed significant main effects for Group ($F(2, 70) = 8.887, p < .001$) Time ($F(1, 910) = 25.575, p < .001$) and Coda Condition ($F(6, 910) = 973.069, p < .001$), and significant interaction effects for Group*Coda Condition ($F(12, 910) = 12.598, p < .001$), Coda Condition*Time ($F(6, 910) = 6.883, p < .001$), and Group*Coda Condition*Time ($F(12, 910) = 1.767, p < .05$), but not for Group*Time ($F(2, 910) = 2.182, p = .113$). These results indicate that there was a significant difference between groups overall (main effect of Group), between coda conditions overall (main effect of Coda Condition), between overall accuracy at Time 1 and Time 2 (main effect of Time), between the groups according to coda condition (interaction of Group*Coda Condition), between coda conditions according to Time (interaction of Coda Condition*Time), and between groups according to Coda Condition and Time (interaction of Group*Coda Condition*Time). However, the groups were not significantly different *overall* according to Time (Group*Time interaction not significant). Since the comparison of the groups in each coda condition over time is the primary focus of the study, the three-way interaction between Group*Coda Condition*Time will be explored.

The primary reason for this interaction effect is that there were no significant differences between the three L2 groups at either data collection time for all coda conditions except for the Asp condition, for which the two study abroad groups made significant positive gains over time while the At-home group did not. Since the three-way interaction of Group*Coda Condition*Time was significant ($p = 0.049$), the data was split by Group and three models were run looking for effects of Time, Coda Condition and the interaction of Time*Coda Condition for each group, with Subject as a random effect. This was to examine the significance of changes over time within each L2 group according to Coda Condition. Table 42 displays a combined summary of the Time 1 and Time 2 data according to Coda Condition for the three L2 learner groups.

Table 42. Identification accuracy for each L2 group by coda condition and time (T1/T2) (SA = study abroad, AH = At-home)

Group		No Coda	Asp	[s]	[r]	[l]	[m]/[n]	[f]
AH	T1 Mean	93.0	0.30	99.4	90.2	96.4	99.6	79.2
	<i>T1 SD</i>	6.80	1.54	1.98	18.24	4.89	2.00	18.01
	T2 Mean	95.9	0.9	99.4	90.7	97.2	99.2	86.0
	<i>T2 SD</i>	6.19	3.38	1.98	16.25	4.58	2.77	17.56
SA1	T1 Mean	94.7	18.1	99.1	93.7	95.2	99.6	81.3
	<i>T1 SD</i>	8.24	17.91	3.27	9.95	8.98	2.09	20.07
	T2 Mean	94.6	35.8	100.0	95.7	98.3	99.6	87.0
	<i>T2 SD</i>	7.94	27.63	0.00	8.03	3.88	2.09	14.60
SA2	T1 Mean	96.4	16.7	100.0	93.3	96.0	99.6	75.2
	<i>T1 SD</i>	4.88	21.73	0.00	11.57	6.46	6.32	21.82
	T2 Mean	97.2	38.2	100.0	93.3	98.4	100.0	86.0
	<i>T2 SD</i>	4.72	30.46	0.00	9.07	3.74	0.00	12.58

Study Abroad 1 Group. The results of the linear mixed effects model analysis for the Study Abroad 1 group showed significant main effects for Coda Condition ($F(1,286) = 211.794, p < .001$) and Time ($F(1,286) = 8.889, p < .01$) along with a significant interaction effect for Coda Condition*Time ($F(1,286) = 2.937, p < .01$). Post-hoc pairwise comparisons using the Sidak correction were conducted to compare the accuracy rates for this group at Time 1 and Time 2 in each coda condition. The mean increase of 17.7 percent from Time 1 to Time 2 in the Asp condition was highly significant ($p < .001$) for this group. The Study Abroad 1 group's accuracy did not change significantly over time for any other coda conditions ($p = .117$ to 1.00).

Study Abroad 2 group. The results for the Study Abroad 2 group are similar in most respects to those of Study Abroad 1. There were significant main effects for Coda Condition ($F(1,312) = 221.943, p < .001$) and Time ($F(1,312) = 14.815, p < .001$) along with a significant interaction effect for Coda Condition*Time ($F(1,312) = 5.288, p < .001$). Post-hoc pairwise comparisons using the Sidak correction were used to compare the accuracy rates for the Study Abroad 2 group at Time 1 and Time 2 in each coda condition. The increase in mean accuracy in the Asp condition was 21.5 percent from Time 1 to

Time 2 and this increase was highly significant ($p < .001$). A result that differed in comparison to the Study Abroad 1 group was that the Study Abroad 2 group's accuracy in the coda [f] condition was significantly higher at Time 2 compared to Time 1 ($p < .01$). Accuracy in all other coda conditions, though, did not change significantly over time ($p = .496$ to 1.00).

At-home group. The results of the model for the At-home group were very different from those of the two study abroad groups. Like the study abroad groups, there was a significant main effect for Coda Condition ($F(1,312) = 750.875, p = <.001$), but no main effect for Time ($F(1,312) = 2.685, p = .102$). Also, there was no significant interaction effect for Coda Condition*Time ($F(1,312) = .947, p = .462$). Post-hoc pairwise comparisons using the Sidak correction were conducted to compare the accuracy rates for the At-home group at Time 1 and Time 2 in each coda condition to show why there was no significant interaction effect for Coda Condition*Time. While both of the study abroad groups increased accuracy in the Asp condition to a significant degree at Time 2, the At-home group did not (Time 1: $M = 0.3\%$, $SD = 1.54$; Time 2: $M = 0.9\%$, $SD = 3.38$), resulting in no significant change over time ($p = .814$). Additionally, there was no significant change over time for most of the other coda conditions ($p = .243$ to 1.00), with the exception of [f]. There was a significant increase of 6.8 percent at Time 2 for this coda condition ($p < .05$).

Summary. The results of the linear mixed model showed that there was significant change over time in the Asp condition for the two study abroad groups, which were exposed to Andalusian /s/-aspiration, while there was no significant change over time for the At-home group, which was not exposed to Andalusian /s/-aspiration. Finally, there were no significant differences for all three groups between the accuracy at Time 1 and Time 2 for the No Coda, [s], [l], [r], and nasal coda conditions. At the same time, there was a significant increase over time for the accuracy in the [f] condition among learners in the Study Abroad 2 and At-home groups, while there was no significant change in this

condition for the Study Abroad 1 group. Figure 9 displays the Time 1 and Time 2 results for accuracy in the Asp condition according to Group.

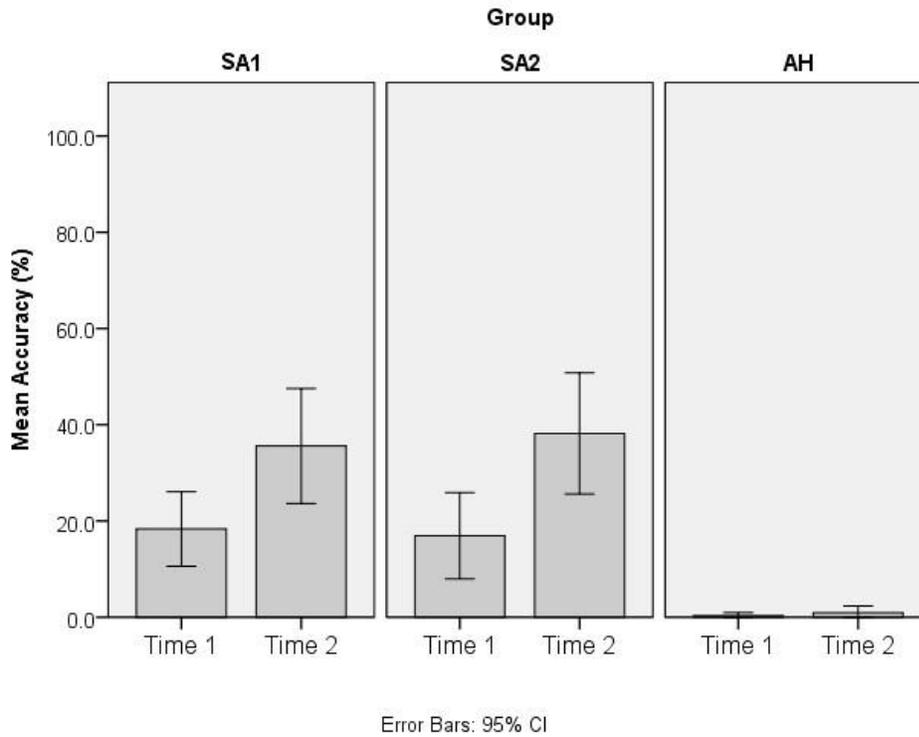


Figure 9. Accuracy in the Asp condition of the identification task over time (L2 groups)

Individual Patterns of the Study Abroad participants in the Aspirated Condition

Since both study abroad groups increased accuracy in the Asp condition over time, an analysis was undertaken to determine how many participants in each of the study abroad groups increased accuracy, stayed the same, or decreased accuracy in the Asp condition from Time 1 to Time 2. This was accomplished by comparing the percent of change from Time 1 to Time 2 for each participant, which was calculated by subtracting the Time 1 mean from the Time 2 mean (T2-T1) for the Asp condition. Table 43 summarizes the findings. Both groups were nearly identical in terms of the distribution of learners who increased accuracy, decreased accuracy, or did not change over time. Additionally, both groups' mean percent increase among the individual learners was very similar. The Study Abroad 2

group, though, showed a slightly wider range of percent positive change, from 1 percent to 78 percent, while the Study Abroad 1 learners' range was between 1 percent and 49 percent. Also, learners in the Study Abroad 1 group who decreased accuracy over time had a wider range of decrease than those in the Study Abroad 2 group, ranging from 11 percent to 33 percent decrease, while the Study Abroad 2 learners ranged from 1 percent to 20 percent decrease. These results, when considered alongside the results of the linear mixed model, show that there was no significant effect of different study abroad programs, as the two study abroad groups were not significantly different from one another at Time 1 or Time 2 in any condition. Importantly, though, the study abroad groups were significantly different from the At-home group in the aspirated condition, indicating that exposure to /s/-aspiration during study abroad played a significant role in accuracy at Time 2.

Table 43. Individual patterns of according to study abroad group for the Asp condition over time. (SA = study abroad)

Group	Accuracy increase N = number of people (mean % increase, SD)	Accuracy decrease N = number of people (mean % decrease, SD)	No change over time N = number of people
SA1 (N=23)	17 (+27.40%, SD = 16.09)	5 (-19.80%, SD = 8.56)	1
SA2 (N=25)	19 (+29.58%, SD = 22.99)	4 (-7.50%, SD = 8.96)	2

Analysis of Response Types by Coda Condition for Each Group

Next, in order to determine the distribution of responses to each coda condition for the NS groups and each L2 learner group at both Time 1 and Time 2, the percentage of each orthographic coda responses (i.e., <s>, <f>, <l>, <r>, nasal, No Coda) given by each group for each coda condition was tabulated by means of a cross-tabulation. The response type analysis will report the distribution of the responses for each of the coda conditions in terms of percentage and total number of responses. First, the response types of the two NS groups are presented, followed by the L2 groups at Time 1 and Time 2.

SEVILLE group. The distribution of response types for the SEVILLE group can be observed in Table 44. In all but the Asp coda condition, the response was in accord with the coda condition at least

96.7 percent of the time. All of the coda conditions exhibited some variation in response types. For the CodaS, [f], [l], [r], nasal and No Coda conditions, the variation was limited to no more than six responses of a given type apart from the correct response. For the Asp condition, however, the responses exhibited more variation.

Table 44. Distribution of Identification responses for the SEVILLE group according to coda condition (N = 33)

Coda Condition	Response						Total
	s	f	l	r	m/n	No Coda	
Asp	74.5% (334)	20.8% (93)		0.2% (1)		4.4% (20)	448
CodaS	97.4% (450)	0.2% (1)	0.2% (1)		1.3% (6)	0.9% (4)	462
[f]	0.6% (2)	97.6% (322)	0.6% (2)		0.3% (1)	0.9% (3)	330
[l]		0.3% (1)	98.2% (324)	0.3% (1)	1.2% (4)		330
[r]	1.3% (4)	0.3% (1)	0.3% (1)	97.3% (289)	0.7% (2)		297
nasal		0.6% (2)	0.6% (2)		98.2% (324)	0.3% (1)	329
No Coda	0.5% (2)	0.9% (4)	0.5% (2)	0.7% (3)	0.7% (3)	96.7% (410)	424

In the Asp condition, the overwhelming tendency of the SEVILLE group was to perceive the aspiration as <s> (74.5%). Additionally, aspiration was perceived as containing a coda [f] 20.8 percent of the time and as having No Coda consonant 4.7 percent of the time. All other responses were limited to no more than five cases per response type.

Non-Aspirating group. The response types for the Non-Aspirating group exhibited less variation than the SEVILLE group. The Non-Aspirating group's responses for the CodaS, [f], [l], [r], nasal and No Coda conditions were all nearly categorical or categorical. However, the Asp condition showed more variation than any other coda condition. The Non-Aspirating group only perceived the aspiration as <s>

26.1 percent of the time, while the predominant response was No Coda (55.2%) and <f> received 17.9 percent of the responses.

Table 45. Distribution of Identification responses for the Non-Aspirating group according to coda condition (N = 10)

		Response					No Coda	Total
		s	f	l	r	m/n		
Coda Condition	Asp	26.1% (35)	17.9% (24)	0.7% (1)			55.2% (74)	134
	CodaS	100.0% (140)						140
	[f]	3.0% (1)	96.0% (96)		1.0% (1)			100
	[l]			100.0% (100)				100
	[r]				98.9% (89)	1.1% (1)		90
	nasal					100.0% (110)		110
	no coda						100% (132)	132

At-home group (Time 1). The At-home group's response types at Time 2 show a fairly consistent pattern for all coda conditions except for the Asp condition and the [f] condition. The CodaS, [l], [r], nasal and No Coda conditions all had at least 90.2 percent accuracy and the other responses were limited to no more than 9.3 percent of the responses. On the other hand, the At-home group responded almost categorically to the Asp items as if there were No Coda consonant (95.8%), with a small variety of other responses. Most importantly, there was only one response out of 335 Asp items that was perceived as <s>. The [f] condition, as well, was below 90 percent accuracy, as 19.6 percent of coda [f] items were perceived as <s>. This is not surprising given that both [f] and [s] are voiceless fricatives.

Table 46. Distribution of Identification responses for the At-home group according to coda condition at Time 1 (N = 25)

		Response					No Coda	Total
		s	f	l	r	m/n		
Coda Condition	Asp	0.3% (1)	0.3% (1)	2.1% (7)	0.9% (3)	0.6% (2)	95.8% (321)	335
	CodaS	99.4% (348)					0.6% (2)	350
	[f]	19.6% (49)	79.2% (198)		0.4% (1)		0.8% (2)	250
	[l]			96.4% (241)	3.2% (8)		0.4% (1)	250
	[r]			9.3% (21)	90.2% (203)		0.4% (1)	225
	nasal					99.6% (249)	0.4% (1)	250
	No Coda			2.4% (8)	3.6% (12)	0.9% (3)	93.0% (307)	330

At-home group (Time 2). The Time 2 data for the At-home group is nearly identical to the Time 1 data. There was an increase of 6.8 percent accuracy for the [f] condition, a very slight increase of +2.9 percent for the No Coda condition, +0.8 percent for the [l] condition (97.2 percent), +0.5 percent for the [r] condition, and a very slight decrease of -0.8 percent for the nasal condition. The CodaS condition was identical to Time 1. Interestingly, though, there was almost no increase in accuracy for the Asp condition, which was almost zero at Time 1. The At-home group a mere +0.6 percent more accurate at Time 2 compared to Time 1 (0.9%). As at Time 1, the Asp condition was perceived as No Coda at a very high rate of 94.4 percent.

Table 47. Distribution of Identification responses for the At-home group according to coda condition at Time 2 (N = 25)

		Response					No Coda	Total
		s	f	l	r	m/n		
Coda Condition	Asp	0.9% (3)	1.8% (6)	2.1% (7)	0.6% (2)	0.3% (1)	94.4% (321)	340
	CodaS	99.4% (348)					0.6% (2)	350
	[f]	12.4% (31)	86.0% (215)	0.8% (2)	0.4% (1)		0.4% (1)	250
	[l]			97.2% (243)	2.4% (6)		0.4% (1)	250
	[r]		0.4% (1)	8.4% (19)	90.7% (204)		0.4% (1)	225
	nasal	0.4% (1)	0.4% (1)			99.2% (248)		250
	No Coda		0.3% (1)	0.9% (3)	2.2% (7)	0.6% (2)	95.9% (307)	320

Study Abroad 1 group (Time 1). The Study Abroad 1 group followed a similar pattern to the At-home group in terms of the control coda conditions, the CodaS condition, and the No Coda condition. In other words, the Study Abroad 1 group's responses were in accord with the coda condition at least 81.3 percent of the time ([f]), but mostly 93.7 percent and above. There was some variation in responses, primarily for [f], which was perceived as <s> 18.3 percent of the time. However, other responses for all coda conditions except the Asp condition were relatively limited to a small number. Nevertheless, the Asp condition showed the most variation. Aspirated items were perceived as having No Coda consonant 75.1 percent of the time, whereas <s> received only 18.1% of the responses. Other responses to Asp stimuli included <l> (3.2%), <f> (2.9%), <r> (0.6%), and <m>/<n> (0.3%).

Table 48. Distribution of Identification responses for the Study Abroad 1 group according to coda condition at Time 1 (N = 23)

		Response					No Coda	Total
		s	f	l	r	m/n		
Coda Condition	Asp	18.1% (56)	2.9% (9)	3.2% (10)	0.6% (2)	0.3% (1)	75.1% (232)	309
	CodaS	99.1% (319)	0.3% (1)			0.6% (2)		322
	[f]	18.3% (42)	81.3% (187)		0.4% (1)			230
	[l]			95.2% (219)	3.5% (6)	0.9% (2)	0.4% (1)	230
	[r]			6.3% (13)	93.7% (194)			207
	Nasal					99.6% (229)	0.4% (1)	230
	No Coda			2.0% (6)	2.3% (7)	1.0% (3)	94.7% (286)	302

Study Abroad 1 group (Time 2). The responses for the Study Abroad 1 group at Time 2 were similar to Time 1 for all coda conditions except for the Asp condition. There was an increase in responses that were in accord with the correct response for the CodaS, [f], [l], and [r] conditions. The nasal condition remained the same and the No Coda condition exhibited a decrease of -0.1 percent at Time 2. Nonetheless, the Study Abroad 1 group associated the Asp items more with <s> (35.8%) at Time 2 than Time 1 (18.1%), and it was the largest increase in accuracy (+17.7%) among all coda conditions. The Study Abroad 1 group also decreased the percentage of No Coda responses by 17.6 percent, indicating that the shift in perception occurred via decreasing associations of Asp items with No Coda and increasing the association of aspiration with <s>. All other responses for the Asp condition remained approximately the same from Time 1 to Time 2.

Table 49. Distribution of Identification responses for the Study Abroad 1 group according to coda condition at Time 2 (N = 23)

	Response						Total
	s	f	l	r	m/n	No Coda	
Asp	35.8% (112)	4.2% (13)	2.2% (7)	0.3% (1)		57.5% (180)	313
CodaS	100.0% (322)						322
[f]	13.0% (30)	87.0% (200)					230
[l]			98.3% (226)		1.7% (4)		230
[r]			3.4% (7)	95.7% (198)		1.0% (2)	207
nasal					99.6% (229)	0.4% (1)	230
No Coda	3.1% (9)		1.4% (4)	1.0% (3)		94.6% (286)	294

Study Abroad 2 group (Time 1). The data for the Study Abroad 2 group's response types at Time 1 are presented in Table 50. At Time 1, the Study Abroad 2 group followed a similar pattern to the Study Abroad 1 group at Time 1. The CodaS condition was perceived categorically as <s>. The [l], [r], nasal and No Coda conditions also had very high rates of perceiving these codas according to the correct orthographic mapping (93.3% to 99.6%). The Asp condition, on the other hand, exhibited more variation, with the predominant response being No Coda (77.9%), followed by <s> (16.7%), <f> (4.5%), <l> (0.6%) and <r> (0.3%).

Table 50. Distribution of Identification responses for the Study Abroad 2 group according to coda condition at Time 1 (N = 25)

		Response					No Coda	Total
		s	f	l	r	m/n		
Coda Condition	Asp	16.7% (56)	4.5% (15)	0.6% (2)	0.3% (1)		77.9% (261)	335
	CodaS	100.0% (350)						350
	[f]	24.8% (62)	75.2% (188)					250
	[l]			96.0% (240)		4.0% (10)		250
	[r]			6.7% (15)	93.3% (210)			225
	nasal			0.4% (1)		99.6% (249)		250
	No Coda	0.9% (3)		0.6% (2)	2.1% (7)		96.4% (286)	330

Study Abroad 2 group (Time 2). Table 51 displays the response types for the Study Abroad 2 group at Time 2 according to Coda Condition. Like the Study Abroad 1 group at Time 2, the perception of the CodaS, [l], [r], nasal, and No Coda conditions increased in accuracy slightly or stayed the same. For these coda conditions, the Study Abroad 2 group perceived in accord with the correct response between 93.3 percent and 100 percent of the time. In addition, there was very little variation. Also like the Study Abroad 1 group at Time 2, the Study Abroad 2 group's accuracy in the [f] condition increased (75.2% to 86.0%). Finally, the Asp condition again showed the most variation in response types and there was an increase in the identification of /s/-aspiration as <s> from Time 1 to Time 2 (+21.5%). This increase brought the Study Abroad 2 group to a slightly higher accuracy rate than the Study Abroad 1 group at Time 2. This was in tandem with a decrease in the perception of Asp items as No Coda (-20.0%). The other response types for the Asp coda condition were limited to a small percentage of the items.

Table 51. Distribution of Identification responses for the Study Abroad 2 group according to coda condition at Time 2 (N = 25)

		Response					No Coda	Total
		s	f	l	r	m/n		
Coda Condition	Asp	38.2% (130)	2.6% (9)	0.3% (1)	0.6% (2)	0.3% (1)	57.9% (197)	340
	CodaS	100.0% (350)						350
	[f]	14.0% (35)	86.0% (215)					250
	[l]			98.4% (246)	0.8% (2)		0.8% (2)	250
	[r]	0.4% (1)		6.2% (14)	93.3% (210)			225
	nasal					100.0% (250)		250
	No Coda	1.6% (5)		0.3% (1)	0.6% (2)	0.3% (1)	97.2% (311)	320

Summary. In all, the response types for all groups for the CodaS, [l], [r], nasal, No Coda, and [f] conditions were highly accurate, demonstrating response types that were in accord with the coda condition at least 75 percent of the time. This figure is misleading, though, because all responses to all coda conditions but the Asp and [f] conditions were in correct at least 90 percent of the time. The [f] condition exhibited more variation in responses, which were primarily <s>. But the coda condition with the most variation in response types within each group and among all the groups was the Asp condition. The native speaker data for the response types given according to each coda condition revealed that the SEVILLE group showed the highest perception of <s> when Asp stimuli were heard (74.5%). The most frequent other response to the Asp stimuli for this group was <f> (20.8%). Additionally, the SEVILLE group perceived No Coda very infrequently when Asp stimuli were heard (4.4%). This is in contrast to the other four groups, which perceived Asp stimuli as No Coda the majority of the time. The Non-Aspirating group identified /s/-aspiration as No Coda 55.2 percent of the time, <f> 17.9 percent of the time, and <s> 26.1 percent of the time. The At-home group identified /s/-aspiration as No Coda 95.5

percent of the time at Time 1 and 94.4 percent of the time at Time 2. The Study Abroad 1 group identified /s/-aspiration as No Coda 75.1 percent of the time at Time 1 and 57.5 percent of the time at Time 2. And finally, the Study Abroad 2 group identified it as No Coda 77.9 percent of the time at Time 1 and 57.9 percent of the time at Time 2. Among the L2 groups, only the two study abroad groups reduced their identification of Asp stimuli as No Coda and increased their identification of /s/-aspiration as <s>. The At-home group did evidence a very small change, but the increase in correct responses to Asp stimuli was only 0.6 percent by Time 2. Table 52 displays the distribution of responses to the Asp stimuli for all groups, including Time 1 and Time 2 L2 data.

Table 52. Distribution of Identification responses to Asp stimuli for all groups, including Time 1 and Time 2 responses for the L2 groups (SA = study abroad, AH = At-home) (N = 116)

	Response						Total
	s	f	l	r	m/n	No Coda	
AH T1	0.3% (1)	0.3% (1)	2.1% (7)	0.9% (3)	0.6% (2)	95.8% (321)	335
AH T2	0.9% (3)	1.8% (6)	2.1% (7)	0.6% (2)	0.3% (1)	94.4% (321)	340
SA1 T1	18.1% (56)	2.9% (9)	3.2% (10)	0.6% (2)	0.3% (1)	75.1% (232)	309
SA1 T2	35.8% (112)	4.2% (13)	2.2% (7)	0.3% (1)		57.5% (180)	313
SA2 T1	16.7% (56)	4.5% (15)	0.6% (2)	0.3% (1)		77.9% (261)	335
SA2 T2	38.2% (130)	2.6% (9)	0.3% (1)	0.6% (2)	0.3% (1)	57.9% (197)	340
SEVILLE	74.5% (334)	20.8% (93)		0.2% (1)		4.4% (20)	448
NON-ASP	26.1% (35)	17.9% (24)	0.7% (1)			55.2% (74)	134

The Effects of Extralinguistic Factors on Study Abroad Learners' Identification Accuracy.

The final analysis of the forced-choice identification task compares the two study abroad groups and the effect that the extralinguistic factors of *NS Contact*, *Language Use* and *Grammar Test Score* on

the study abroad learner group's accuracy over time in the aspirated test condition. Importantly, the At-home group was not included in this analysis. The primary reason is that there was no change over time for this group in the Asp condition and an extremely low accuracy rate at Time 1 and Time 2, and therefore there is no possibility to show any sort of effect for any of the extralinguistic factors on the At-home group's accuracy over time in the perception of /s/-aspiration. The second reason is that the purpose of this analysis is to answer the third research question, which seeks to determine the effects of extralinguistic factors on *study abroad learners'* perception of a dialect-specific phonological variant to which they are exposed over time.

To investigate the relationships between accuracy and the independent variables, a bivariate correlation analysis was conducted. The data was split by Coda Condition so as to focus the analysis of extralinguistic factors on the perception of the Asp condition since there was little to no significant change over time in the other coda conditions.

Language Use. The first extralinguistic factor that was analyzed to determine its relationship to increased accuracy at Time 2 in the perceptual identification of /s/-aspiration was the learners' use of Spanish during the semester. To remind the reader, this was operationalized in a few different ways. First, learners were asked at the end of their study abroad to report the overall percentage of Spanish use (compared to English) during the semester on a scale of one to 100. This was also done for the use of Spanish specifically for reading, writing and speaking separately. Then, the learners were also asked to report the number of hours out of an average 16-hour day that they spent doing the following things: speaking with native speakers of Spanish in Spanish and in English, speaking with native speakers of English in Spanish and English, watching media in Spanish and English, and reading in Spanish and English. Since some learners erroneously reported more hours than the 16 hour maximum outlined in the instructions, the raw number of hours was converted to a percentage of the total reported time.

Descriptives. Before proceeding to the results of the correlation analysis it is necessary to present the descriptive statistics for the language use factors for the study abroad groups. Table 53 presents the mean, standard deviation, minimum, and maximum for the responses provided by the study abroad learners on the questionnaire at Time 2.¹⁶

¹⁶ Statistics broken down into the two groups of study abroad learners can be found in Appendix N.

Table 53. Descriptive statistics for the language use factors as reported by the study abroad learners on the questionnaire at Time 2 (Identification task)

	Factors	Mean	SD	Min	Max
	% use of Spanish overall during the semester	57.50%	14.90%	30.00%	90.00%
	% use of Spanish when writing during the semester	52.80%	20.00%	10.00%	93.00%
	% use of Spanish when reading during the semester	53.90%	23.30%	10.00%	91.00%
	% use of Spanish speaking	56.30%	16.60%	24.00%	90.00%
Both study abroad groups combined	% of avg. day speaking Spanish with NS of Spanish	13.10%	6.80%	3.00%	31.00%
	% of avg. day speaking Spanish with NS of English	9.30%	7.40%	0.00%	27.00%
	% of avg. day speaking English with NS of Spanish	1.50%	3.10%	0.00%	13.00%
	% of avg. day speaking English with NS of English	14.90%	7.40%	1.00%	31.00%
	% of avg. day spent watching media in Spanish	4.50%	5.60%	0.00%	31.00%
	% of avg. day spent watching media in English	4.90%	4.50%	0.00%	18.00%
	% of avg. day spent reading in Spanish	7.50%	5.70%	0.00%	24.00%
	% of avg. day spent reading in English	5.40%	3.60%	0.00%	16.00%

In order to determine whether any of these language use factors was significantly correlated with accuracy in the Asp condition at Time 2 as well as the differential between accuracy at Time 1 and Time 2 for the two study abroad groups, a bivariate correlation analysis using Spearman's rank correlation coefficient¹⁷ was conducted (Table 54). Of the language use factors analyzed, only the percent use of Spanish overall ($\rho = .285, p < .05$), percent use of Spanish while writing ($\rho = .417, p < .01$) and the amount of time during an average day speaking in Spanish with NSs of Spanish ($\rho = .377, p < .01$) were significantly and positively correlated with Time 2 accuracy in the Asp condition. There were a few other factors that were approaching significance: The percent use of Spanish while speaking ($\rho = .251, p = .085$), time spent speaking in Spanish with NSs of English ($\rho = .255, p = .081$), and a negative correlation between watching various media in English and accuracy in the Asp condition at Time 2 ($\rho = -.250, p = .087$).

Table 54. Spearman's rank correlations for accuracy in the aspirated condition at Time 2 and language use factors.¹⁸

Factors	Spearman's Rank Correlation Coefficient (ρ)	p value
% use of Spanish overall	.285	.05*
% use of Spanish - writing	.417	.003**
% use of Spanish - reading	.228	.118
% use of Spanish - speaking	.251	.085
Time spent speaking with NSs of Spanish in Spanish	.377	.008**
Time spent speaking with NSs of Spanish in English	-.207	.158
Time spent speaking with NSs of English in Spanish	.255	.081
Time spent speaking with NSs of English in English	-.070	.637
Time spent watching media in Spanish	-.013	.929
Time spent watching media in English	-.250	.087
Time spent reading in Spanish	.003	.984
Time spent reading in English	.016	.914
** Correlation significant at $p < .01$ * Correlation significant at $p < .05$		

¹⁷ Spearman's correlation coefficient was used because it is more resistant than Pearson's correlation coefficient to data that is not normally distributed. The data for the extralinguistic factors tended to not be normally distributed.

¹⁸ See Appendix L for scatterplots of the significant correlations of each task

Correlations among Language Use Factors

It is important to note that some of the language use factors were significantly correlated with one another. For example, the reported percent use of Spanish overall during the semester was positively correlated with the reported percent use of Spanish when speaking ($\rho = .747, p < .001$) and was negatively correlated with the amount of time spent on an average day speaking English with NSs of English ($\rho = -.284, p = .05$). This shows that those who reported using more Spanish overall also reported more use of Spanish specifically when speaking and reported speaking less English with NSs of English. Likewise, the reported amount of time spent on an average day speaking Spanish with NSs of Spanish was positively correlated with the reported percent use of Spanish (speaking) ($\rho = .320, p < .05$) and the reported amount of time on an average day spent speaking with English speakers in Spanish ($\rho = .433, p < .01$). Thus, those who spent more time speaking Spanish with native speakers of Spanish also reported speaking more Spanish with English speakers. Finally, there were positive correlations between the time spent speaking in Spanish with NSs of Spanish and the percent use of Spanish when writing ($\rho = .148, p < .01$), and between the percent use of Spanish overall and the percent use of Spanish when writing ($\rho = .176, p < .001$).

Time 2-Time 1 Accuracy Differential and Language Use Factors

Next, a bivariate correlation analysis was run to determine whether the language use factors were correlated with the percent change over time in the mean accuracy rate in the Asp condition. The accuracy differential was calculated by subtracting the Time 1 mean from the Time 2 mean in the Asp condition for each participant, resulting in some positive and some negative values (i.e., when a participant's accuracy decreased at Time 2). The Study Abroad 1 group showed a mean increase of 17.2 percent ($SD = 24.9$) and the Study Abroad 2 group showed a mean increase of 21.3 percent ($SD = 24.9$). The accuracy differential was positively correlated with Time 2 accuracy in the Asp condition, as would

be expected ($\rho = .744, p < .001$). There was a positive correlation between the Time 2-Time 1 accuracy differential and the reported time spent on the average day speaking Spanish with NSs of Spanish ($\rho = .300, p < .05$), as well as the percent use of Spanish when writing ($\rho = .320, p < .05$). All other correlations were not significant (see Table 55). Thus, those whose accuracy in the Asp condition increased from Time 1 to Time 2 also reported using more Spanish when writing and spending more time speaking in Spanish with NSs of Spanish on an average day.

Table 55. Spearman's rank correlations for Time1 to Time 2 accuracy change over time in the aspirated condition and language use factors.

Factors	Spearman's Rank Correlation Coefficient (ρ)	<i>p</i> value
% use of Spanish overall	.228	.119
% use of Spanish - writing	.320	.027*
% use of Spanish - reading	.242	.098
% use of Spanish - speaking	.179	.222
Time spent speaking with NSs of Spanish in Spanish	.300	.039*
Time spent speaking with NSs of Spanish in English	.001	.996
Time spent speaking with NSs of English in Spanish	.035	.815
Time spent speaking with NSs of English in English	-.040	.787
Time spent watching media in Spanish	.097	.514
Time spent watching media in English	-.235	.108
Time spent reading in Spanish	.122	.410
Time spent reading in English	.019	.900
** Correlation significant at $p < .01$ * Correlation significant at $p < .05$		

Summary. Of the language use factors that were analyzed to determine their relationship to accuracy in the Asp condition both at Time 2 and the differential between Time 1 and Time 2, only the overall reported percent use of Spanish during the semester, the percent use of Spanish when writing, and the amount of time that learners spent speaking Spanish with NSs of Spanish on an average day were positively correlated with Time 2 accuracy in the Asp condition and/or the Time 2-Time 1 accuracy differential. The results of these correlations indicate that those learners who spent more time speaking

Spanish with NSs, using Spanish when writing, and using more Spanish overall during the semester also had higher identification accuracy at Time 2 for Asp stimuli and greater positive change in accuracy between Time 1 and Time 2.

Native Speaker Contact. The second extralinguistic factor to be analyzed was the amount of contact the study abroad learners reported having with native Spanish speakers during their semester abroad. NS contact and language use are most often interconnected, since it is impossible to use Spanish when speaking to NSs of Spanish without having contact with them. Thus, the language use factor of the time spent on an average day speaking in Spanish to native Spanish speakers, which was significantly correlated with both Time 2 accuracy and Time Time 2-Time 1 accuracy differential, also reflects contact with NSs of Spanish in the target dialect regions. Nevertheless, NS contact was also operationalized as two variables: the number of native speakers with whom the learners had conversations of 15 to 20 minutes or more in Spanish on a weekly basis (WeeklyContact) and also on a daily (DailyContact) basis. The range of responses allowed on the Time 2 questionnaire was zero to 12 for the number of NSs that the participants could choose for each of these two variables. The mean number of NSs with whom the study abroad learners (i.e., both groups combined) reported having contact on a weekly basis was 4.7 ($SD = 2.6$, $Min = 1$, $Max = 12$), and on a daily basis was 2.8 ($SD = 1.9$, $Min = 0$, $Max = 8$)¹⁹. Both variables were tested for correlation with both the Time 2 mean accuracy in the Asp coda condition and the Time 2-Time 1 accuracy differential using a bivariate correlation analysis with Spearman's rank correlation coefficient.

WeeklyContact did not correlate significantly with the Time 2-Time 1 accuracy differential ($\rho = .009$, $p = .951$) or Time 2 accuracy in the Asp condition ($\rho = .046$, $p = .754$). The DailyContact variable, too, did not correlate significantly with the Time 2 – Time 1 accuracy differential ($\rho = .055$, $p = .711$) or

¹⁹ A table with the breakdown of the two study abroad groups' reports of ContactWeekly and ContactDaily can be found in Appendix N.

Time 2 accuracy ($\rho = .231, p = .176$). Both ContactWeekly and ContactDaily, though, were positively correlated with percent use of Spanish overall ($\rho = .288, p < .05$) and ($\rho = .319, p < .05$) respectively. This shows that there was at least some consistency between those who reported using more Spanish during the semester and those who had contact with a greater number of NSs of Spanish. However, despite this, the number of NSs of Spanish with whom the learners had contact did not significantly correlate with their performance on the identification task for the Asp condition.

Grammar Test Score. The final extralinguistic factor of interest is the study abroad learners' grammar test score at Time 1 and Time 2. The first question related to grammar test score is whether the learners' score at Time 1 was related to their initial accuracy in the Asp coda condition at Time 1. First, an independent samples *t*-test was run to compare the two study abroad groups' mean grammar test scores at Time 1. The Study Abroad 1 group's mean score was 13.04 ($SD = 3.69$) out of a possible 20 points and the Study Abroad 2 group's mean score was 10.96 ($SD = 3.69$). The *t*-test showed that the Study Abroad 1 group's Time 1 grammar score was significantly higher than that of Study Abroad 2 ($t(670) = 7.31, p < .001$). In order to test for a correlation between an individual's Time 1 grammar score and mean accuracy in the Asp coda condition, a bivariate correlation analysis was conducted. Using Spearman's rank correlation coefficient, there was no significant correlation between an individual's T1 grammar score and the mean accuracy at Time 1 in the Asp condition for the study abroad learners ($\rho = .171, p = .244$).

The second question related to grammar test score is whether there was a correlation between the Time 2 grammar test score and Time 2 accuracy in the Asp condition. At Time 2, the Study Abroad 1 group's mean grammar test score was 16.09 ($SD = 3.26$) and the Study Abroad 2 group's mean score was 13.16 ($SD = 3.59$) and the Study Abroad 1 group's score was again significantly higher than the Study Abroad 2 group's score ($t(670) = 11.03, p < .001$). The correlation analysis showed that Time 2 grammar score and Time 2 accuracy in the Asp condition were not significantly correlated ($\rho = .139, p = .346$).

The final question related to grammar score is whether the change over time in grammar score for each participant was correlated with the change over time in accuracy in the Asp condition. To accomplish this, the mean percent change (positive or negative) in accuracy from Time 1 to Time 2 in the Asp condition was calculated for each participant. The change in grammar test score was calculated by subtracting the Time 1 score from the Time 2 score. This was then tested for correlation against the Time 2 – Time 1 accuracy differential. There was no significant correlation ($\rho = .074$, $p = .616$).

Grammar Test Score Summary. The results of the analysis of the effect of grammar test score on identification accuracy in the Asp coda condition revealed no significant correlations between study abroad learners' grammar test scores at both data collection times and their accuracy in the Asp coda condition at both data collection times for the study abroad learners. Additionally, the differences between the Time 1 and Time 2 grammar test scores and between the Time 1 and Time 2 mean accuracy rates in the Asp condition were not significantly correlated. These results taken together show no significant effect of grammar test score on identification accuracy in the Asp coda condition for the study abroad learners.

Results of the Lexical Decision Task

Now that the results of the forced-choice identification have been described in detail, the results of the lexical decision task will be presented. The analysis of the lexical decision task consisted of two parts: the analysis of accuracy (i.e., correctly accepting real words or correctly rejecting non-words) and the analysis of response time. First, the analysis of accuracy will be presented, comparing the factors of Group (SEVILLE, Non-Aspirating, Study Abroad 1, Study Abroad 2, At-home) and Condition (CodaS real words, CodaS non-words, Asp real words, Asp non-words, Distracter real words, Distracter non-words) at Time 1, followed by Time 2, and then a comparison of the three L2 learner groups according to the factor Time (i.e., Time 1 compared to Time 2). This will be followed by an analysis of the

relationships between the extralinguistic factors of NS Contact, Language Use, and Grammar Score. Then, the analysis of the response times from the lexical decision task will follow.

Lexical Decision Accuracy for each Condition according to Group (Time 1)

In order to compare the effects of Group and Condition on lexical decision accuracy at Time 1, the Time 2 L2 learner data were excluded and a linear mixed effects model was run that included Subject as a random effect and Group and Condition as fixed effects. The dependent variable was the calculated mean percent correct for each participant in each condition. Post-hoc pairwise comparisons are reported using the Sidak correction. The model showed significant main effects for Group ($F(4, 111) = 72.264, p < .001$) and Condition ($F(5, 555) = 123.088, p < .001$), as well as a significant interaction effect for Group*Condition ($F(20, 555) = 15.518, p < .001$). The significant main effect for Group occurred because there were significant differences overall among the groups at Time 1. The SEVILLE group was significantly more accurate overall (i.e., regardless of condition) than all three L2 groups ($p < .001$) but was not significantly different from the Non-Aspirating group ($p = .920$). The Non-Aspirating group was also more accurate overall than the three L2 groups ($p < .001$). The two study abroad groups were not significantly different from one another overall ($p = 1.00$), but were both significantly more accurate than the At-home group ($p < .001$).

The significant main effect of Condition came about because of significant differences found among the overall accuracy rates of all groups combined in the different conditions. Though various differences were found, the most striking are that the Asp real word condition showed significantly lower accuracy than all other conditions ($p < .001$) and the CodaS real word and Distracter real word conditions were not significantly different ($p = 1.00$), but both showed significantly higher accuracy rates than most of the other conditions (generally $p < .001$). This shows an overall difference of how the

participants responded to real words with /s/-aspiration compared to real words that had a coda sibilant (CodaS) or distracters that had neither aspiration or a sibilant.

The interaction between Group and Condition was significant because there was a significant effect of Condition for each of the groups ($p < .01$ to $.001$) except for the SEVILLE group ($p = .074$), according to the univariate tests. This interaction will be explored further. First, accuracy in each of the Conditions will be compared according to Group using Post-hoc pairwise comparisons with the Sidak correction. Table 56 displays the mean accuracy data for each Condition according to Group.

Table 56. Mean percent accuracy of lexical decision responses for each Condition by Group for NSs and L2 learners at Time 1 (parentheses denote standard deviation)

Condition	Group				
	SEVILLE	NON-ASP	SA1	SA2	AH
CodaS real words	96.28 (4.49)	95.84 (6.46)	93.36 (6.60)	90.32 (6.70)	84.57 (10.41)
CodaS non-words	94.21 (6.75)	92.73 (6.14)	69.96 (18.63)	65.82 (23.48)	61.45 (16.40)
Asp real words	90.23 (10.75)	76.61 (20.60)	53.70 (17.84)	52.47 (24.10)	22.02 (16.84)
Asp non- words	94.09 (6.78)	92.50 (4.86)	82.17 (15.58)	85.40 (12.49)	83.40 (12.56)
DIS real words	97.83 (2.41)	97.92 (2.43)	90.98 (6.07)	92.98 (4.61)	87.55 (7.12)
DIS non- words	91.38 (7.19)	90.98 (6.21)	70.16 (11.49)	71.14 (13.86)	62.43 (16.06)

CodaS Real Word Condition (Time 1). In the CodaS real word condition at Time 1 there was only one significant group difference. No groups were significantly different ($p = .137$ to 1.00) except that the SEVILLE group ($M = 96.28\%$, $SD = 4.49$) was significantly more accurate than the At-home group ($M = 84.57\%$, $SD = 10.41$, $p < .01$).

CodaS Non-word Condition (Time 1). In the CodaS non-word condition, the SEVILLE ($M = 94.21\%$, $SD = 6.75$) and Non-Aspirating ($M = 92.73\%$, $SD = 6.14$) groups were not significantly different ($p = 1.00$), while they were both significantly more accurate than the Study Abroad 1 ($M = 69.96\%$, $SD = 18.63$), Study Abroad 2 ($M = 65.82\%$, $SD = 23.48$), and At-home ($M = 61.45\%$, $SD = 16.40$) groups at the $p < .001$ level. The three L2 groups were not significantly different from one another ($p = .168$ to $.943$).

Asp Real Word Condition (Time 1). The Asp real word condition showed the lowest overall mean of 58.97 percent for all groups combined. In this condition, interestingly, the SEVILLE group ($M = 90.23\%$, $SD = 10.75$) was significantly more accurate than all other groups (L2 groups: $p < .001$; Non-Aspirating: $p < .05$). Additionally, the Non-Aspirating group ($M = 76.61\%$, $SD = 20.60$), though it was significantly less accurate than the SEVILLE group, was significantly more accurate than the three L2 groups ($p < .001$). The Study Abroad 1 ($M = 53.70\%$, $SD = 17.84$) and Study Abroad 2 ($M = 52.47\%$, $SD = 24.10$) groups were not significantly different ($p = 1.00$) and both were significantly more accurate than the At-home group ($M = 22.02\%$, $SD = 16.84$, $p < .001$).

Asp Non-word Condition (Time 1). In the Asp non-word condition at Time 1, the SEVILLE group ($M = 94.09\%$, $SD = 6.78$) was not significantly different from the Non-Aspirating group ($M = 92.50\%$, $SD = 4.86$, $p = 1.00$) and the Study Abroad 2 group ($M = 85.40\%$, $SD = 12.49$, $p = .083$), though the latter comparison was approaching significance. Conversely, the SEVILLE group was significantly more accurate than the Study Abroad 1 ($M = 82.17\%$, $SD = 15.58$, $p < .01$) and At-home groups ($M = 83.40\%$, $SD = 12.56$, $p < .05$). The Non-Aspirating group was not significantly different from any other group ($p = .253$ to 1.00) and the three L2 learner groups were not significantly different from one another ($p = .990$ to 1.00).

Distracter Real Word Condition (Time 1). For the Distracter real word condition, the SEVILLE group ($M = 97.83\%$, $SD = 2.41$) was significantly more accurate than the At-home group ($M = 87.55\%$, $SD = 7.12$, $p < .05$), but was not different from any other group ($p = .356$ to 1.00). The Non-Aspirating group ($M = 97.92\%$, $SD = 2.43$), again, was not significantly different from any other group ($p = .233$ to 1.00). Finally, the Study Abroad 1 ($M = 90.98\%$, $SD = 6.07$), Study Abroad 2 ($M = 92.98\%$, $SD = 4.61$) and At-home ($M = 87.55\%$, $SD = 7.12$) groups were not significantly different from one another in the Distracter real word condition ($p = .731$ to 1.00).

Distracter Non-word Condition (Time 1). Finally, in the Distracter non-word condition at Time 1, the SEVILLE group ($M = 91.38\%$, $SD = 7.19$) was significantly more accurate than all three L2 groups ($p <$

.001). The Non-Aspirating group ($M = 90.98\%$, $SD = 6.21$), which was not significantly different from the SEVILLE group, was more accurate than all three L2 groups ($p < .001$). The Study Abroad 1 ($M = 70.16\%$, $SD = 11.49$), Study Abroad 2 ($M = 71.14\%$, $SD = 13.86$), and At-home ($M = 62.43\%$, $SD = 16.06$) groups were not significantly different in the Distracter non-word condition ($p = .128$ to 1.00).

Summary. The results of the linear mixed effects model showed that the SEVILLE and Non-Aspirating groups were not significantly different in their mean accuracy for all conditions except the Asp real word condition, in which the SEVILLE group was more accurate than the Non-Aspirating group. Likewise, the three L2 groups were not significantly different from one another in all conditions except the Asp real word condition, where the two study abroad groups were significantly more accurate than the At-home group at Time 1. Both NS groups were significantly more accurate in the Asp real word condition than the three L2 groups. The Non-Aspirating group was intermediate between the L2 groups and the SEVILLE group in the Asp real word condition, as it was significantly less accurate than the SEVILLE group but significantly more accurate than the three L2 groups. In other conditions though, the Non-Aspirating group was sometimes significantly more accurate than the L2 groups (e.g., CodaS non-words, Distracter non-words) and other times not (e.g., CodaS real words, Asp non-words, Distracter real words). Overall, the SEVILLE group was more accurate than the L2 groups to varying degrees depending on Condition.

Lexical Decision Accuracy for each Group according to Condition (Time 1)

Next, a comparison of each group's accuracy in the different conditions will be presented separately based on the post-hoc pairwise comparisons from the linear mixed effects model using the Sidak correction.

SEVILLE group. The SEVILLE group's accuracy across the different conditions was consistent, as there were no significant differences among the conditions ($p = .111$ to 1.00).

Non-Aspirating group. The Non-Aspirating group was similar to the SEVILLE group overall in that the accuracy across most of the conditions was not significantly different ($p = .949$ to 1.00), with the only exception being that accuracy in the Asp real word condition was significantly lower than the CodaS real word ($p < .01$), CodaS non-word ($p < .05$), Asp non-word ($p < .05$), and Distracter real word ($p < .01$) conditions, but not the Distracter non-word condition ($p = .081$), though it was approaching significance.

Study Abroad 1 group (Time 1). The Study Abroad 1 group's accuracy across conditions was not nearly as categorical as the two NS groups. Accuracy in the CodaS real word condition was not different from that in the Distracter real word condition ($p = 1.00$), but was significantly higher compared to accuracy in all other conditions (Asp non-word: $p < .05$; all others: $p < .001$). Accuracy in the Asp real word condition was lower than all other conditions ($p < .001$). The CodaS non-word condition accuracy was higher than that of the Asp real word condition ($p < .001$) and lower than the CodaS real word ($p < .001$), Asp non-word ($p < .01$) and Distracter real word ($p < .001$) conditions, but was not different from the Distracter non-word condition ($p = 1.00$). The Asp non-word condition had higher accuracy than the CodaS non-word ($p < .01$), Distracter non-word ($p < .01$) and Asp real word ($p < .001$) conditions, but was not different from the Distracter real word condition ($p = .141$). The accuracy in the Distracter real word condition was higher than the CodaS non-word, Asp real word and Distracter non-word conditions ($p < .001$).

Study Abroad 2 group (Time 1). The Study Abroad 2 group followed the same basic pattern as the Study Abroad 1 group at Time 1 according to Condition. The CodaS real word condition exhibited significantly higher accuracy than all other conditions ($p < .001$) except the Asp non-word ($p = .883$) and Distracter real word ($p = 1.00$) conditions. The lack of significance for the CodaS real word and Asp non-word comparison was different from the Study Abroad 1 group. Additionally, the Distracter real word and Asp non-word conditions were not significantly different ($p = .271$). The Asp real word condition exhibited accuracy that was significantly lower than in all other conditions (CodaS non-word: $p < .01$; all

others: $p < .001$). Finally, the accuracy in the CodaS non-word and Distracter non-word conditions was not different ($p = .809$).

At-home group (Time 1). At Time 1, the At-home group was similar to the two study abroad groups in terms of the comparisons among the conditions. Like the Study Abroad 2 group, accuracy in the CodaS real word condition was significantly higher than in the Asp real word, CodaS non-word and Distracter non-word conditions ($p < .001$), but was not different from the Asp non-word and Distracter real word conditions ($p = .999$ to 1.00), which both also exhibited higher accuracy than all other conditions ($p < .001$). The Asp real word condition showed significantly lower accuracy compared to all other conditions ($p < .001$). The CodaS non-word and Distracter non-word conditions were not different ($p = 1.00$).

Summary. The analysis of each group's accuracy according to Condition revealed important differences between the groups. First, the SEVILLE group and the Non-Aspirating group were very similar and exhibited very high accuracy across all conditions, though the Non-Aspirating group's accuracy in the Asp real word condition was significantly lower than in all other conditions (except the Distracter non-word condition) while the SEVILLE group did not differ across conditions. The L2 groups all showed a similar pattern according to Condition in that each group was significantly more accurate in the CodaS real word, Distracter real word and Asp non-word conditions than the other three conditions. Importantly, the L2 groups and Non-Aspirating group found it quite easy to reject Asp non-words, while at the same time finding it difficult to accept Asp real words. The SEVILLE group, as expected, did not make such a distinction between Asp real word and non-words.

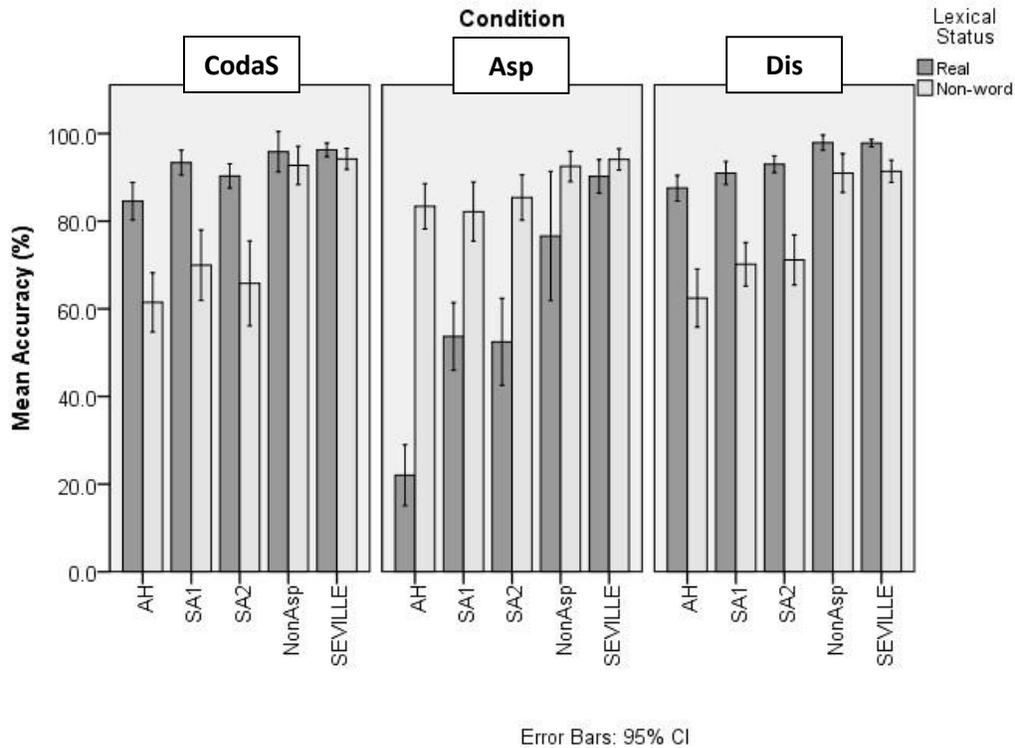


Figure 10. Summary of Mean lexical decision accuracy (%) for each Condition according to Group and Lexical Status (Time 1)

Lexical Decision Accuracy for each Condition according to Group (Time 2)

In order to compare the interaction of Group and Condition at Time 2, the Time 1 L2 learner data was excluded and a linear mixed effects model was run that included Subject as a random effect and Group and Condition as fixed effects. The dependent variable was the calculated mean percent accuracy for each participant in each condition. Post-hoc pairwise comparisons were conducted using the Sidak correction. The model showed significant main effects for Group ($F(1,111) = 63.547, p < .001$) and Condition ($F(1,555) = 91.280, p < .001$), as well as a significant interaction effect for Group*Condition ($F(1,555) = 21.134, p < .001$). The significant main effect of Group showed that The SEVILLE and Non-Aspiring groups' accuracy rates overall were significantly higher than those of the L2 groups ($p < .001$) and were not significantly different from one another ($p = .899$). The two study abroad groups' accuracy rates overall were not significantly different ($p = 1.00$) and were significantly higher

than that of the At-home group ($p < .001$). The significant main effect of Condition showed that, as at Time 1, the accuracy rate of all groups combined was lower in the Asp real word condition than all other conditions ($p < .001$). Also, the Distracter real word condition and CodaS real word condition were again not significantly different ($p = .982$), but did show higher accuracy than all other conditions ($p < .001$) except for the Asp non-word condition ($p = .514$ & $.982$) respectively. Finally, the interaction of Group*Condition was significant and will be explored further. First, accuracy in each of the Conditions will be compared according to Group using Post-hoc pairwise comparisons. Table 52 displays the mean accuracy data for each Condition according to Group.

Table 57. Mean percent accuracy of lexical decision responses for each Condition by Group at Time 2 (parentheses denote standard deviation)

Condition	SEVILLE	NON-ASP	SA1	SA2	AH
CodaS real words	96.28 (4.49)	95.84 (6.46)	92.45 (7.41)	92.74 (6.80)	80.98 (12.56)
CodaS non- words	94.21 (6.75)	92.73 (6.14)	83.00 (12.28)	81.45 (17.10)	71.45 (19.71)
Asp real words	90.23 (10.75)	76.61 (20.60)	69.38 (20.89)	72.71 (19.80)	14.85 (15.03)
Asp non-words	94.09 (6.78)	92.50 (4.86)	90.00 (10.45)	86.00 (15.88)	89.40 (16.03)
DIS real words	97.83 (2.41)	97.92 (2.43)	93.52 (6.99)	94.72 (4.87)	83.62 (8.74)
DIS non-words	91.38 (7.19)	90.98 (6.21)	82.18 (9.68)	79.37 (12.04)	68.24 (15.54)

CodaS Real Word Condition (Time 2). At Time 2, the only significant difference in the CodaS real word condition was that the At-home group ($M = 80.98\%$, $SD = 12.56$) was significantly less accurate than all other groups (SEVILLE: $p < .001$; all others: $p < .01$). The Study Abroad 1 ($M = 92.45\%$, $SD = 7.41$), Study Abroad 2 ($M = 92.74\%$, $SD = 6.80$), SEVILLE ($M = 96.28\%$, $SD = 12.49$) and Non-Aspirating ($M = 95.84\%$, $SD = 6.46$) groups were not significantly different in the CodaS real word condition ($p = .932$ to 1.00).

CodaS Non-word Condition (Time 2). In the CodaS non-word condition, the accuracy of the SEVILLE group ($M = 94.21\%$, $SD = 6.75$, $p < .01$) was significantly higher than that of all L2 groups for Time 2 data (AH: $p < .001$; Study Abroad 1/Study Abroad 2: $p < .01$). The At-home group ($M = 71.45\%$, $SD =$

19.71) was again significantly less accurate than all other groups (Study Abroad 1: $p < .01$; Study Abroad 2: $p < .05$; NS groups: $p < .001$). The Study Abroad 1 group ($M = 83.00\%$, $SD = 12.28$) was not different from the Study Abroad 2 ($M = 81.45\%$, $SD = 17.10$, $p = 1.00$) and Non-Aspirating ($p = .270$) groups, but was less accurate than the SEVILLE group and more accurate than the At-home group ($p < .01$). The Study Abroad 2 group followed the exact same pattern as the Study Abroad 1 group.

Asp Real Word Condition (Time 2). In the Asp real word condition at Time 2, the L2 learner groups were significantly less accurate at $p < .001$ than the SEVILLE group ($M = 90.23\%$, $SD = 10.75$). Contrary to Time 1, both Study Abroad 1 ($M = 69.38\%$, $SD = 20.89$) and Study Abroad 2 ($M = 72.71\%$, $SD = 19.80$) were not different than the Non-Aspirating group ($M = 76.61\%$, $SD = 20.60$, $p = .684$ to $.992$) at Time 2, and both study abroad groups were significantly more accurate than the At-home group ($M = 14.85$, $SD = 15.03$ $p < .001$). The At-home group was significantly less accurate than all other groups ($p < .001$).

Asp Non-word Condition (Time 2). In the Asp non-word condition at Time 2 for the L2 learners, there were no significant differences between any groups ($p = .099$ to 1.00). However, the difference between the Study Abroad 2 group ($M = 86.00\%$, $SD = 15.88$) and the SEVILLE group ($M = 94.09\%$, $SD = 6.78$) was approaching significance ($p = .099$), as the Study Abroad 2 group's accuracy was the lowest in this condition.

Distracter Real Word Condition (Time 2). In the Distracter real word condition at Time 2, the At-home group ($M = 83.62\%$, $SD = 8.74$) was significantly less accurate than the SEVILLE ($M = 97.83\%$, $SD = 2.41$, $p < .001$), Non-Aspirating ($M = 97.92\%$, $SD = 2.43$, $p < .05$), Study Abroad 1 ($M = 93.52\%$, $SD = 6.99$, $p < .05$) and Study Abroad 2 ($M = 94.72\%$, $SD = 4.87$, $p < .05$) groups. Nevertheless, none of the other groups were significantly different ($p = .866$ to 1.00).

Distracter Non-Word Condition (Time 2). Finally, in the Distracter non-word condition, the At-home group ($M = 68.24\%$, $SD = 15.54$) was again significantly less accurate than all other groups (Study

Abroad 1: $p < .01$; Study Abroad 2: $p < .05$; NS groups: $p < .001$). The Study Abroad 1 group ($M = 82.18\%$, $SD = 9.68$) was not different from the Study Abroad 2 group ($M = 79.37\%$, $SD = 12.04$, $p = .995$) and Non-Aspirating group ($M = 90.98\%$, $SD = 6.21$, $p = .407$), but was significantly less accurate than the SEVILLE group ($M = 91.38\%$, $SD = 7.19$, $p < .05$). The Study Abroad 2 group followed the same pattern, but was nearing significance in being less accurate than the Non-Aspirating group ($p = .088$).

Summary. Across the board, the At-home group's accuracy was significantly lower than all other groups in all conditions except for Asp non-words, where all groups were statistically equal. The two study abroad groups and the NS groups were also equal in the CodaS real word and Distracter real word conditions, while the At-home group lagged behind. The greatest difference among the L2 groups was in the Asp real word condition, where the Study Abroad 1 and Study Abroad 2 groups increased accuracy at Time 2, becoming no different than the Non-Aspirating group but still being lower in accuracy than the SEVILLE group. In the Asp real word condition, there was a decrease at Time 2 for the At-home group, and thus this group's accuracy was lower by a large margin than that of all other groups. Interestingly, the two study abroad groups, even then they were not statistically equivalent to the NS groups, patterned alongside one another at Time 2, as they were never statistically different.

Lexical Decision Accuracy for each Group according to Condition (Time 2)

Next, the results for Group according to Condition will be presented separately using the post-hoc pairwise comparisons with the Sidak correction. Only the L2 results at Time 2 will be described as to not repeat the NS results unnecessarily. However, NS results will be presented in a summary figure that compares all groups.

Study Abroad 1 group (Time 2). At Time 2 for the Study Abroad 1 group, the CodaS real word, Asp non-word and Distracter real word conditions were not significantly different ($p = 1.00$) and the accuracy in the CodaS non-word condition was nearly significant in being lower than that of CodaS real

words ($p = .055$). The Asp real word condition showed accuracy that was significantly below that of Distracter non-word ($p < .01$) and all other conditions ($p < .001$).

Study Abroad 2 group (Time 2). The Study Abroad 2 group at Time 2 follows the same type of pattern as the Study Abroad 1 group. The CodaS real word, Asp non-word and Distracter real word conditions were not significantly different ($p = .376$ to 1.00) and the accuracy in the CodaS non-word condition was significantly lower than that of CodaS real word ($p < .01$). Accuracy in the Distracter non-word condition was not significantly different from that of the CodaS non-word ($p = 1.00$), Asp real word ($p = .393$), and Asp non-word ($p = .402$) conditions, but was significantly lower than accuracy in the CodaS real word and Distracter real word ($p < .001$) conditions. The Asp real word condition showed accuracy that was significantly below that of the CodaS real word ($p < .001$), Distracter real word ($p < .01$) and Asp non-word ($p < .001$) conditions, but was not different from the Distracter non-word ($p = .393$) and CodaS non-word conditions, though the latter was approaching significance ($p = .074$).

At-home group (Time 2). The At-home group's patterns by condition at Time 2 were different in some ways from those of the two study abroad groups. Some results were similar. For example, accuracy in the CodaS real word, Asp non-word and Distracter real word conditions was still not significantly different ($p = .101$ to $.999$) and the accuracy in the CodaS non-word condition was significantly lower than that of the CodaS real word condition ($p < .05$). Also, accuracy for Distracter non-words was not significantly different from CodaS non-words ($p = .995$), but in the case of the At-home group, accuracy for Distracter non-words was significantly higher than that of the Asp real word condition ($p < .001$) and lower than that of Asp non-word ($p < .001$), Distracter real word ($p < .001$) and CodaS real word conditions ($p < .001$). The Asp real word condition displayed accuracy that was significantly below that of all other conditions ($p < .001$).

Summary. The comparisons of each condition within each L2 group at Time 2 showed that the two study abroad groups patterned similarly, while the At-home group differed in some ways. All three

groups' accuracy was lowest in the Asp real word condition compared to other conditions to a significant degree. One difference was that the Study Abroad 2 group's accuracy in the Asp real word condition was not significantly different from accuracy in the Distracter non-word condition. Thus, the study abroad groups' Asp real word accuracy was beginning to approach that of other easier conditions at Time 2. All three groups were similar in that accuracy in the CodaS real word, Asp non-word and Distracter real word conditions was not significantly different. Nevertheless, where the At-home group departs from the other two is that the Asp non-word and Distracter real word conditions are reversed compared to the study abroad groups. Figure 11 displays a summary of the results for each L2 group at Time 2 according to Condition.

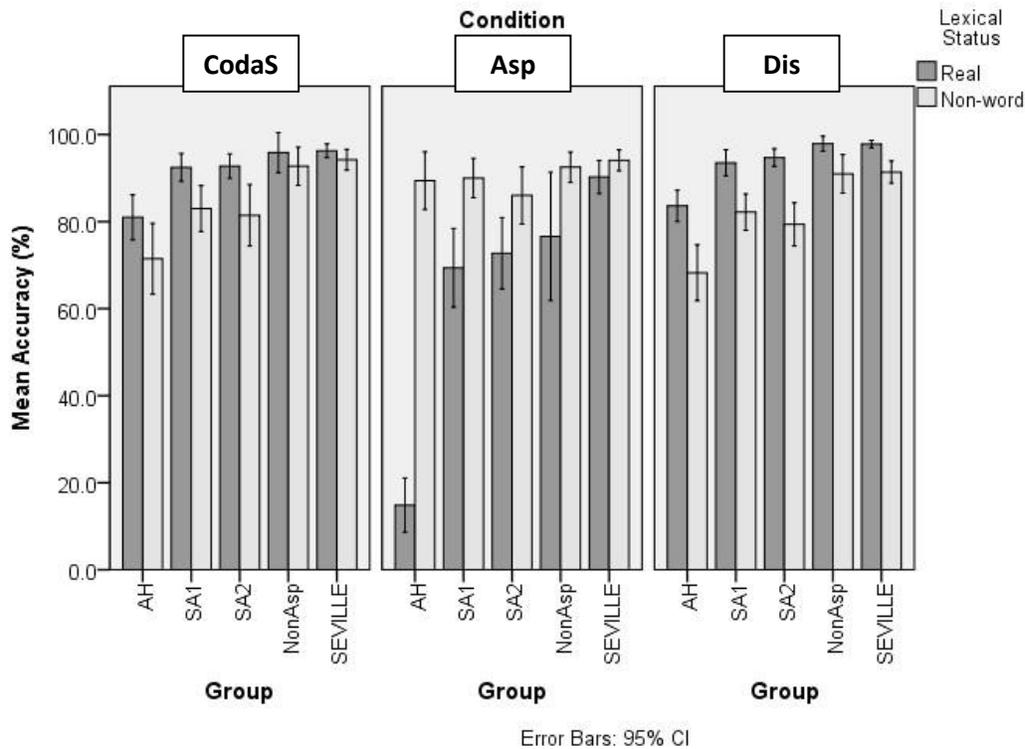


Figure 11. Summary of Mean Lexical decision Accuracy for each Group according to Condition and Lexical Status (Time 2 – Note: NS data is the same as at Time 1)

Analysis of the Effect of Time on the L2 Learners' Lexical Decision Patterns

Now that the results for Condition according to Group and for Group according to Condition have been presented for Time 1 and Time 2 separately, the L2 learner groups will be compared over time in order to determine whether the Study Abroad 1 and Study Abroad 2 groups' exposure to Andalusian /s/-aspiration led to a difference between the study abroad and at-home groups. In order to test the variable Time, both Time 1 and Time 2 data for the L2 learners (excluding NSs) were analyzed using a linear mixed effects model in which the mean percent accuracy for each participant in each condition was the dependent variable, Subject was a random effect, and Group, Condition and Time were fixed effects. The model revealed significant main effects for Group ($F(1,770) = 29.664, p < .001$), Condition ($F(1,770) = 231.937, p < .001$) and Time ($F(1,770) = 45.752, p < .001$), as well as significant interaction effects for Condition*Group ($F(1,770) = 23.750, p < .001$), Condition*Time ($F(1,770) = 6.501, p < .001$), Group*Time ($F(1,770) = 7.403, p < .01$) and Condition*Group*Time ($F(1,770) = 2.524, p < .01$). The main effect of Group showed that the two study abroad groups were overall nearly identical while the At-home group's accuracy was lower. The main effect of Condition was driven by differences in accuracy according to Condition, primarily that the CodaS real word condition, Distracter real word condition, and Aspirated non-word condition showed very high accuracy (86.1% - 90.6%), while the CodaS non-word and Distracter non-word conditions showed intermediate accuracy (72%), and the Aspirated real word condition showed the lowest accuracy (47.5%). The main effect of Time was driven by an overall increase in accuracy from Time 1 to Time 2 (73.3% - 79.2%). In order to explore the most important interaction effect, Condition*Group*Time, post-hoc comparisons with the Sidak correction will be presented in order to compare the accuracy for each group separately according to Time in each Condition, as well as a comparison between the groups according to Time and Condition. As a summary, Table 58 presents the L2 groups' mean percent accuracy and standard deviations in each condition at Time 1 and Time 2.

Table 58. Mean percent accuracy for each L2 group in each condition at Time 1 and Time 2 (parentheses denote standard deviation)

	SA1	SA2	AH
Condition*Time			
CodaS real words Time 1	93.36 (6.60)	90.32 (6.70)	84.57 (10.41)
CodaS real words Time 2	92.45 (7.41)	92.74 (6.80)	80.98 (12.56)
CodaS non words Time 1	69.96 (18.63)	65.82 (23.48)	61.45 (16.40)
CodaS non words Time 2	83.00 (12.28)	81.45 (17.10)	71.45 (19.71)
Asp real words Time 1	53.70 (17.84)	52.47 (24.10)	22.02 (16.84)
Asp real words Time 2	69.38 (20.89)	72.71 (19.80)	14.85 (15.03)
Asp non words Time 1	82.17 (15.58)	85.40 (12.49)	83.40 (12.56)
Asp non words Time 2	90.00 (10.45)	86.00 (15.88)	89.40 (16.03)
DIS real words Time 1	90.98 (6.07)	92.98 (4.61)	87.55 (7.12)
DIS real words Time 2	93.52 (6.99)	94.72 (4.87)	83.62 (8.74)
DIS non words Time 1	70.16 (11.49)	71.14 (13.86)	62.43 (16.06)
DIS non words Time 2	82.18 (9.68)	79.37 (12.04)	68.24 (15.54)

Study Abroad 1 group (Time 1 vs. Time 2). Comparing Time 1 and Time 2 results for the Study Abroad 1 group in each condition reveals that there was a significant increase in accuracy in the CodaS non-word ($p < .01$), Asp real word ($p < .001$), Asp non-word ($p < .05$) and Distracter non-word ($p < .01$) conditions, but that there was no significant change over time in the CodaS real word ($p = .810$) and Distracter real word ($p = .504$) conditions for this group. Figure 12 presents the results of the Study Abroad 1 group at Time 1 and Time 2 according to Condition.

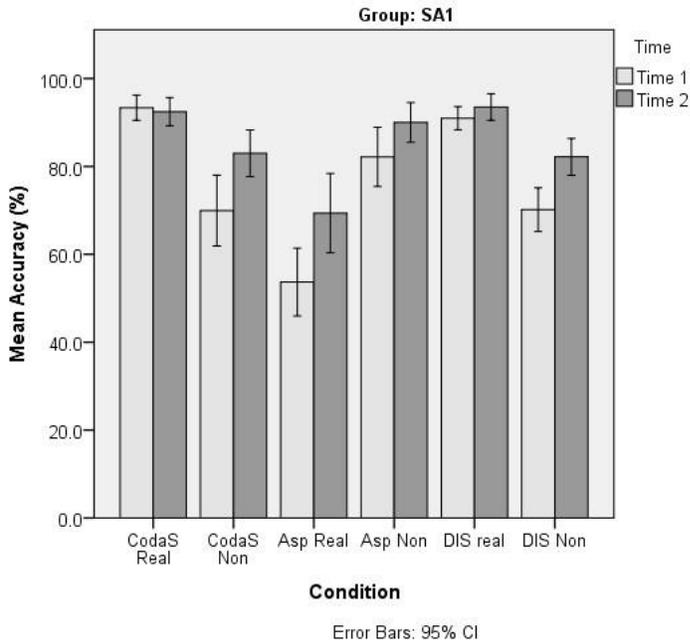


Figure 12. Mean lexical decision accuracy by Condition and Time for the Study Abroad 1 group (N=23)

Study Abroad 2 group (Time 1 vs. Time 2). Comparing the Study Abroad 2 group’s mean accuracy in each condition at Time 1 and Time 2 reveals that this group showed a significant increase in accuracy at Time 2 for the CodaS non-word ($p < .001$), Asp real word ($p < .001$) and Distracter non-word ($p < .05$) conditions, but that there was no significant change for the CodaS real word ($p = .507$), Asp non-word ($p = .869$) and Distracter real word ($p = .634$) conditions. Figure 13 displays the accuracy results on the lexical decision task for the Study Abroad 2 group at Time 1 compared to Time 2 in each condition.

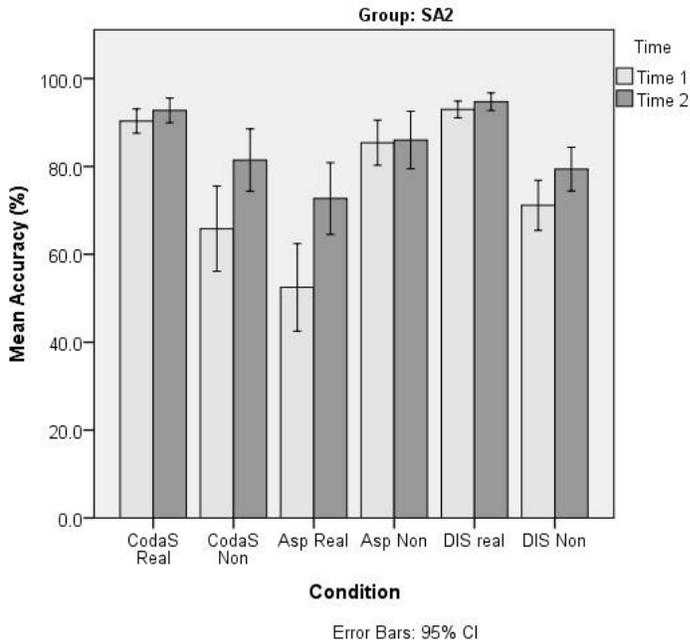


Figure 13. Mean lexical decision accuracy by Condition and Time for the Study Abroad 2 group (N=25)

At-home group (Time 1 vs. Time 2). The At-home group showed a significant increase in accuracy from Time 1 to Time 2 in the CodaS non-word condition ($p < .01$), but not in any other condition. The trend of the At-home group was to decrease in accuracy in the CodaS real word, Distracter real word and Asp real word conditions and increase accuracy in the CodaS non-word, Asp non-word and Distracter non-word conditions. Nevertheless, the change over time was not significant in the CodaS real word ($p = .325$), Asp non-word ($p = .100$), Distracter real word ($p = .282$) and Distracter non-word ($p = .112$) conditions. Strikingly, though, the At-home group decreased in accuracy by 7.17 percent in the Asp real word condition from Time 1 to Time 2, and this change was significant ($p = .05$). Figure 14 displays the results for the At-home group at Time 1 and Time 2 according to Condition.

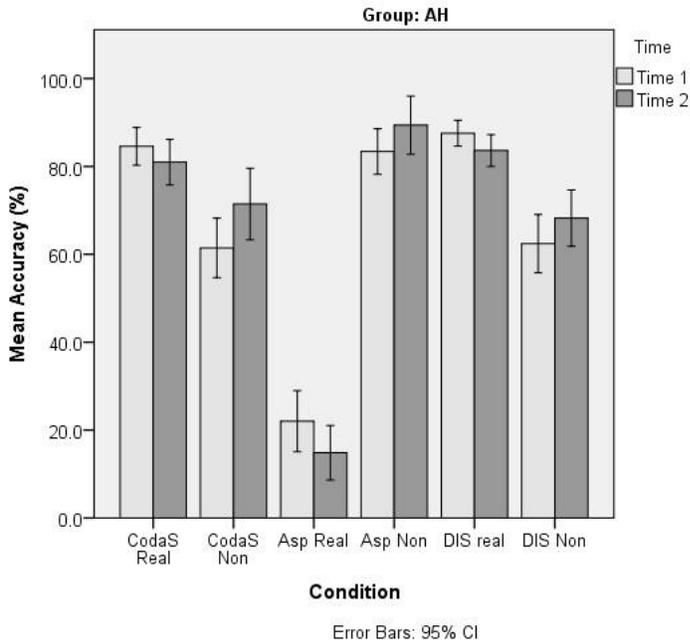


Figure 14. Mean lexical decision accuracy by Condition and Time for the At-home group (N=25)

Overall summary of the L2 groups at Time 1 versus Time 2. Over time, the two study abroad groups patterned very similarly. Their trend in each condition was to increase accuracy at Time 2, and the increase was significant for three (Study Abroad 2) to four (Study Abroad 1) out of the six conditions for these two groups. Post-hoc comparisons of the Study Abroad 1 and Study Abroad 2 groups in each condition at each time showed that accuracy was not significantly different between the two study abroad groups at Time 1 or Time 2 in any condition ($p = .697$ to 1.00). Importantly, both of the study abroad groups' accuracy increased significantly in the Asp real word condition over time and the groups were not different in this condition. Thus, it can be concluded that, as for the forced-choice identification, study abroad program was not a significant factor in determining the study abroad learners' accuracy over time in the aspirated condition for the lexical decision task, or any other condition for that matter.

The At-home group showed opposite trends over time compared to the study abroad groups. The At-home group's accuracy was not significantly different from that of the study abroad groups at

Time 1 for the CodaS real word, CodaS non-word, Asp non-word, Distracter real word and Distracter non-word conditions ($p = .088$ to $.987$). At Time 2, though, the At-home group was significantly less accurate than both study abroad groups in the CodaS real word ($p < .05$), CodaS non-word ($p < .05$), Distracter real word ($p < .05$) and Distracter non-word conditions (Study Abroad 1: $p < .01$; Study Abroad 2: $p < .05$). This is in part because, at Time 2, the At-home group's accuracy decreased in three of the six conditions, while the other three conditions showed an upward trend in accuracy. But the increase was only significant in the CodaS condition. Specifically looking at the Asp real word condition, the At-home group was significantly less accurate than both study abroad groups at Time 1 ($p < .001$) and Time 2 ($p < .001$). Importantly, at Time 2, the At-home group's accuracy in the Asp real word condition decreased significantly while the study abroad groups' accuracy in this condition increased significantly. Figure 15 summarizes the accuracy results for the three L2 groups at Time 1 and Time 2 according to Condition.

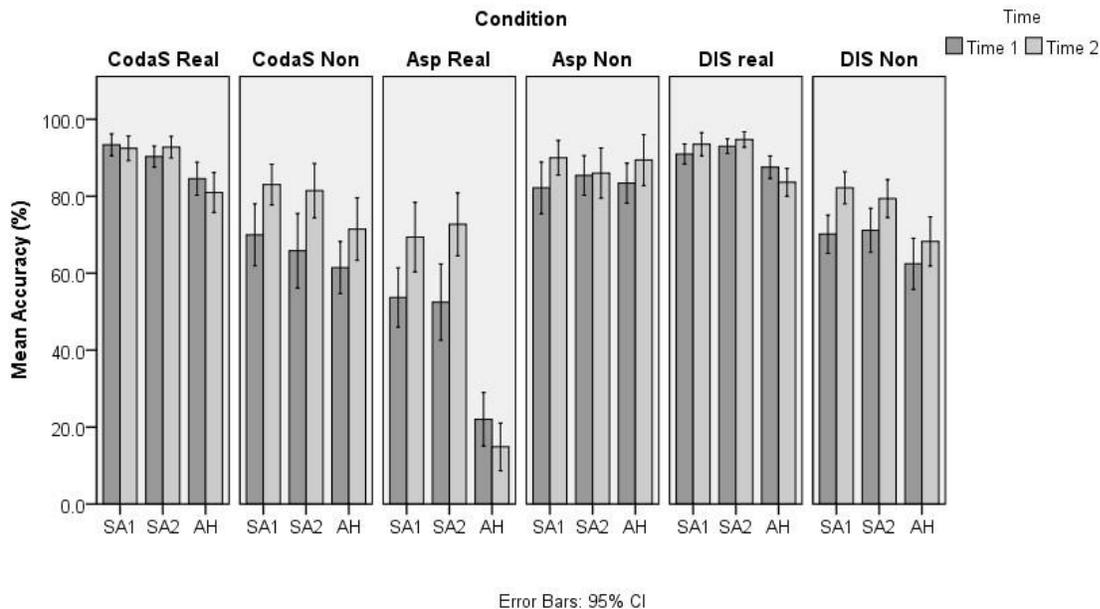


Figure 15. Summary of the lexical decision accuracy results for the L2 groups according to Time and Condition (N=73)

Individual patterns of Study Abroad participants in the Asp real word condition. An analysis was conducted to determine how many participants in each of the study abroad groups increased accuracy,

stayed the same, or decreased accuracy in the aspirated condition from Time 1 to Time 2. This was accomplished by comparing the percent of change from Time 1 to Time 2 for each participant, which was calculated by subtracting the Time 1 mean from the Time 2 mean for the aspirated condition. Table 59 summarizes the findings. For both study abroad groups, there were no learners that did not change in either direction over time. Both groups were similar in terms of the distribution of learners who increased accuracy and decreased accuracy given that the majority of learners in each group increased accuracy in this condition over time. In fact, out of all of the study abroad learners combined (N=48), 39 increased accuracy in the Asp real word condition over time, while only nine decreased accuracy. Additionally, both groups' mean percent increase among the individual learners was very similar. The Study Abroad 2 group, though, showed a slightly wider range of positive change, as demonstrated by a larger standard deviation. The groups were similar in terms of the percent decrease, though the study abroad group's mean percent decrease was slightly larger than that of the Study Abroad 2 group.

Table 59. Number of participants in each study abroad group that increased, decreased, and stayed the same in the Asp real word condition from Time 1 to Time 2.

Group	Accuracy increase N = number if participants (mean % increase, SD)	Accuracy decrease N = number of participants (mean % decrease, SD)	No change
SA1 (N=23)	17 (+29.04%, SD = 14.01)	6 (-22.15%, SD = 15.35)	0
SA2 (N=25)	22 (+25.50%, SD = 21.37)	3 (-18.42%, SD = 15.82)	0

Analysis of the Effects of the Extralinguistic Factors on Study Abroad Learners' Lexical Decision Patterns

The final analysis of the lexical decision task compares the two study abroad groups and the effect that the extralinguistic factors of *NS Contact*, *Language Use* and *Grammar Test Score* on the study abroad learner groups' accuracy over time in the aspirated test condition. As for the analysis of the forced-choice identification task, the At-home group was not included in the analysis of extralinguistic factors. The primary reason is that there was no positive change over time for this group in the Asp real

word condition and very low accuracy rates at Time 1 and 2. The second reason is that the purpose of this analysis is to answer the third research question, which seeks to determine the effects of extralinguistic factors on *study abroad learners'* perception of a dialect-specific phonological variant to which they are exposed over time.

To investigate the relationships between accuracy and the extralinguistic factors, a bivariate correlation analysis was conducted. The data was split by Condition so as to focus the analysis of extralinguistic factors on the perception of the Asp real word condition. Correlations are reported using Spearman's Rank Correlation Coefficient.

Language Use

Descriptives. Before proceeding to the results of the correlation analysis it is necessary to present the descriptive statistics for the language use factors for the study abroad groups. Table 53 is repeated as Table 60 to show the mean, standard deviation, minimum, and maximum for the responses provided by the study abroad learners on the questionnaire at Time 2²⁰

According to the correlation analysis, of the language use factors analyzed, only *Time spent speaking with NSs of Spanish in English* ($\rho = -.357, p < .05$), *Time spent watching media in English* ($\rho = -.373, p < .001$) and *Time spent reading in Spanish* ($\rho = .290, p < .05$) were significantly correlated with Time 2 accuracy in the Asp real word condition. The first two, which involve the use of English, were negatively correlated with accuracy, while time spent reading in Spanish was positively correlated. Table 60 displays the correlation results for all language use factors.

²⁰ Statistics broken down into the two groups of study abroad learners can be found in Appendix N.

Table 60. Descriptive statistics for the language use factors as reported by the study abroad learners on the questionnaire at Time 2

		Factors	Mean	SD	Min	Max
Both study abroad groups combined		% use of Spanish overall during the semester	57.5%	14.9%	30.0%	90.0%
		% use of Spanish when writing during the semester	52.8%	20.0%	10.0%	93.0%
		% use of Spanish when reading during the semester	53.9%	23.3%	10.0%	91.0%
		% use of Spanish speaking	56.3%	16.6%	24.0%	90.0%
		% of avg. day speaking Spanish with NS of Spanish	13.1%	6.8%	3.0%	31.0%
		% of avg. day speaking Spanish with NS of English	9.3%	7.4%	0.0%	27.0%
		% of avg. day speaking English with NS of Spanish	1.5%	3.1%	0.0%	13.0%
		% of avg. day speaking English with NS of English	14.9%	7.4%	1.0%	31.0%
		% of avg. day spent watching media in Spanish	4.5%	5.6%	0.0%	31.0%
		% of avg. day spent watching media in English	4.9%	4.5%	0.0%	18.0%
		% of avg. day spent reading in Spanish	7.5%	5.7%	0.0%	24.0%
		% of avg. day spent reading in English	5.4%	3.6%	0.0%	16.0%

Table 61. Spearman's rank correlations for lexical decision accuracy in the Asp real word condition at Time 2 and language use factors.

Factors	Spearman's Rank Correlation Coefficient (ρ)	<i>p</i> value
% use of Spanish overall	.070	.636
% use of Spanish - writing	.179	.224
% use of Spanish - reading	.007	.963
% use of Spanish - speaking	.013	.932
Time spent speaking with NSs of Spanish in Spanish	.189	.198
Time spent speaking with NSs of Spanish in English	-.357	.013*
Time spent speaking with NSs of English in Spanish	.073	.623
Time spent speaking with NSs of English in English	.157	.288
Time spent watching media in Spanish	-.163	.270
Time spent watching media in English	-.373	.009**
Time spent reading in Spanish	.290	.045*
Time spent reading in English	-.026	.862
** Correlation significant at $p < .01$ * Correlation significant at $p < .05$		

Correlations among language use factors. Interestingly, the factors of *time spent reading in Spanish* and *time spent watching media in English* were negatively correlated to a significant degree ($\rho = -.319, p < .05$), showing that those who reported reading more in Spanish also reported watching less media in English. Additionally, those who reported spending more time watching media in English also reported spending less time speaking in Spanish with native speakers of English ($\rho = -.306, p < .05$). These correlations indicate that the questionnaire did pick up on some important differences between the learners in the study abroad context that indicated that some spent more time using Spanish than English in at least some domains. It appears that, even though the use of Spanish overall and when speaking with NSs of Spanish were not positively correlated with accuracy, as they were for the forced-choice identification task, there was at least a negative relationship between some aspects of the use of English with accuracy on the lexical decision task.

Time 2 – Time 1 Accuracy Differential and Language Use Factors. Next, the Time 2-Time 1 accuracy differential was calculated by subtracting the Time 1 accuracy in each condition for each participant from the Time 2 accuracy. The accuracy differential was significantly correlated with Time 2 accuracy, as expected ($\rho = .574, p < .001$). The accuracy differential in the Asp real word condition was then tested for correlation against the same language use factors. Only one factor was significant and was positively correlated with the T2-T1 accuracy differential: Time spent speaking English with native English speakers ($\rho = .319, p < .05$). This indicated that those who made more positive change over time on the lexical decision task in the Asp real word condition also reported speaking more English with English speakers.

Language use summary. Within the domain of language use, there were few factors in the current study that were significantly correlated with Time 2 accuracy and Time 2-Time 1 accuracy differential. There were two factors related to the use of English by the participants that were significantly and negatively correlated with Time 2 accuracy. On the other hand, reading more in Spanish

was significantly and positively correlated with Time 2 accuracy. However, none of these factors was correlated with the Time 2-Time 1 accuracy differential, while the amount of time spent speaking in English with native English speakers was positively correlated with change over time (T2-T1 accuracy differential).

NS Contact. Next, the results of the bivariate correlation analysis of the mean accuracy at Time 2 for the lexical decision task and the two NS contact factors, ContactWeekly and ContactDaily, are presented. Recall that ContactWeekly and ContactDaily were coded as the number of native Spanish speakers with whom a given participant had conversations of 15 to 20 minutes or more on a weekly and daily basis respectively. Based on the questionnaire data, the mean number of NSs with whom the study abroad learners (i.e., both groups combined) reported having contact on a weekly basis was 4.7 ($SD = 2.6$, $Min = 1$, $Max = 12$), and on a daily basis was 2.8 ($SD = 1.9$, $Min = 0$, $Max = 8$)²¹. The bivariate correlation analysis revealed that neither factor was significantly correlated with Time 2 accuracy (ContactWeekly: $\rho = .068$, $p = .646$; ContactDaily: $\rho = -.054$, $p = .715$) or Time 2-Time 1 accuracy differential (ContactWeekly: $\rho = .119$, $p = .240$; ContactDaily: $\rho = -.060$, $p = .684$).

Grammar Test Score. The final extralinguistic factor is the study abroad learners' grammar test scores at Time 1 and Time 2. The first question related to grammar test score is whether the learners' score at Time 1 was related to their initial accuracy in the Asp real word condition, which could indicate a positive effect of having a higher level of language proficiency prior to SA. First, recall that an independent samples *t*-test was run to compare the two study abroad groups' mean grammar test scores at Time 1, finding that the Study Abroad 1 group's mean score was 13.04 ($SD = 3.69$) out of a possible 20 points was significantly higher ($t(670) = 7.31$, $p < .001$) than the Study Abroad 2 group's mean score of 10.96 ($SD = 3.69$). In order to test for a correlation between an individual's Time 1

²¹ A table with the breakdown of the two study abroad groups' reports of ContactWeekly and ContactDaily can be found in Appendix N.

grammar score and mean accuracy in the Asp real word condition, a bivariate correlation analysis was conducted. Using Spearman's rank correlation coefficient, there was no significant correlation between T1 grammar score and the mean accuracy at Time 1 in the Asp real word condition for the study abroad learners ($\rho = .108, p = .465$). Interestingly, though, there was a significant positive correlation between Time 1 grammar score and Time 1 accuracy for all other conditions ($\rho = .285$ to $.465, p = .001$ to $.05$), showing that learners in the study abroad groups who had a higher Time 1 grammar score also had higher accuracy for all conditions except for the Asp real word condition.

The second question related to grammar test score is whether there was a correlation between the Time 2 grammar test score and Time 2 accuracy in the aspirated condition. At Time 2, the Study Abroad 1 group's mean grammar test score was 16.09 ($SD = 3.26$) and the Study Abroad 2 group's mean score was 13.16 ($SD = 3.59$) and the Study Abroad 1 group's score was again significantly higher than the Study Abroad 2 group's score ($t(670) = 11.03, p < .001$). The correlation analysis showed that Time 2 grammar score and Time 2 accuracy in the Asp real word condition were not significantly correlated ($\rho = .107, p = .468$). However, the Time 2 grammar score was significantly correlated with Time 2 accuracy for the CodaS non-word ($\rho = .356, p < .05$) and Distracter non-word conditions ($\rho = .446, p < .01$), and was marginally significant for Distracter real word ($\rho = .276, p = .057$), Asp non-word ($\rho = .246, p = .091$) and CodaS real word ($\rho = .255, p = .08$) conditions. Therefore, at Time 2, a higher score on the grammar test generally correlated (significantly or nearly significantly) with higher accuracy for all conditions except the Asp real word condition.

The final question related to grammar score is whether the change over time in grammar score for each participant was correlated with the change over time in accuracy in the aspirated condition. To accomplish this, the mean percent change (positive or negative) in accuracy from Time 1 to Time 2 in the aspirated condition was calculated for each participant (i.e., T2-T1 accuracy differential). Likewise, the change in grammar test score was calculated by subtracting the Time 1 score from the Time 2 score. This

was then tested for correlation against the T2-T1 accuracy differential. There was no significant correlation ($\rho = -.040, p = .500$).

Grammar Test Score Summary. The results of the analysis of the effect of grammar test score on identification accuracy in the Asp real word condition revealed no significant correlations between study abroad learners' grammar test scores at the two data collection times and their accuracy in the Asp real word condition at the two data collection times for the study abroad learners. Additionally, the differences between the Time 1 and Time 2 grammar test scores and between the Time 1 and Time 2 mean accuracy rates in the aspirated condition were not significantly correlated. Still, there was a positive correlation trend for the grammar test score at each data collection time and the accuracy in all conditions other than the Asp real word condition. Thus, the Asp real word condition was the exception to the rule that those with higher grammar scores also exhibited higher accuracy. These results taken together show no significant effect of grammar test score on identification accuracy in the Asp real word condition for the study abroad learners. In other words, in the current study, there is no link between overall proficiency (i.e., grammatical) and the acquisition of an optional phonological variant.

Analysis of Lexical Decision Response Time

The final analysis to be presented is that of the response times (RT) from the lexical decision task. In order to analyze the RTs, they were first trimmed to exclude all RTs that were below or above twice the standard deviation for each group. In all, 1,409 data points were removed from the analysis, or 4.9 percent of the dataset. Also, only RTs for correct responses were analyzed. Once cleaned, the RTs were transformed using a log₁₀ transformation to create a normal distribution. For the purposes of the statistical model, log(RT) was the dependent variable. Descriptive statistics will also be reported using means and standard deviations based on the original RTs in milliseconds since they are on a normal time scale that is more easily interpretable. The log(RT) data were analyzed by means of multiple linear mixed

effects models in which Subject and Item were random effects and Condition (CodaS real words, CodaS non-words, Asp real words, Asp non-words, Distracter real words, Distracter non-words), Group (SEVILLE, Non-Aspirating, Study Abroad 1, Study Abroad 2, At-home), Number of Syllables (i.e., two or three), and Time (Time 1, Time 2) were the fixed factors. The first model excluded Time and the second included it in order to analyze change over time. Importantly, the analyses were controlled for the number of syllables in the word, as some words consisted of two syllables and others three. A preliminary analysis showed that both versions of the lexical decision task were not different in terms of the distribution of two and three syllable items so that each participant heard the same distribution of each. In order to test for an effect of Version on the RT data, an independent samples *t*-test was conducted with $\log(\text{RT})$ as the dependent variable and Version as the independent variable (i.e., Version 1 vs. Version 2). The results showed a significant effect of Version on $\log(\text{RT})$ ($t(20034.98) = -4.29, p < .001$), namely that the mean $\log(\text{RT})$ for Version 1 ($M = 3.052, SD = 0.096$) was significantly lower than the mean \log response time for Version 2 ($M = 3.057, SD = 0.094$). Given the inherent difference in length between two and three-syllable words, the fact that the distracter conditions consisted of a greater percentage of two-syllable words than the other conditions, and the significant *t*-test result for Version, the factor of Syllables (i.e., number of syllables in the word) was controlled in the linear mixed model analysis. The first analysis compares the RTs in all conditions at Time 1 according to Group and then the RTs of each group separately according to Condition. This is followed by the same analysis of Time 2 data. Finally, the change over time from Time 1 to Time 2 is presented for the L2 learner groups.

Before proceeding further with the RT results, there is an important disclaimer to help the reader in the interpretation of the RT results. Since the statistical model controlled for the number of syllables in each word and the mean RTs (in milliseconds) reported in tables and figures are not controlled for number of syllables, on a few occasions comparisons between mean RTs for groups or conditions that are reported in tables and figures appear to not align with the significance of post-hoc

pairwise comparisons. Perhaps there is a difference of 100 milliseconds between groups in a certain condition that appears to be significant in a figure or table but the p -value says it is not. Or perhaps there are two conditions for which the RTs appear to be the same but the p -value says they are significantly different. This is the exception rather than the rule, but the reason that this occurs occasionally is that differences between the RTs for two- compared to three-syllable words may be greater or lesser for certain conditions than others or for certain groups than others. Therefore, while the reported means in milliseconds are helpful for seeing overall patterns of response times in each condition for each group on a normal time scale (ms), the p -values, based on the dependent variable of syllable-controlled $\log(\text{RT})$, represent the true nature of the data due to accounting for durational differences between two- and three-syllable words. Thus, $\log(\text{RT})$ is reported in separate tables from the regular RT (in milliseconds). The estimated marginal means for $\log(\text{RT})$, which are the means that are controlled for number of syllables in the model, are reported alongside $\log(\text{RT})$. The reader should look to the estimated marginal means reported in $\log(\text{RT})$ tables in parentheses, as well as the p -values of the pairwise comparisons, as the determiners of true similarity and difference, as they are truer representations of the data than the RTs reported using milliseconds.

Results for each Condition according to Group (Time 1)

First, in order to compare the groups' RTs according to Condition at Time 1, the Time 2 data was excluded and a linear mixed effects model was run in which Subject and Item were random effects and Group, Condition and Syllables were fixed effects. The model showed significant main effects for Syllables ($F(1, 349.02) = 140.00, p < .001$), Condition ($F(5, 360.57) = 90.50, p < .001$) and Group ($F(4, 112.72) = 9.92, p < .001$), as well as a significant interaction effect of Condition*Group ($F(20, 15839.11) = 8.12, p < .001$). The main effect of Syllables was significant because the mean $\log(\text{RT})$ for two-syllable words was significantly lower than that of three-syllable words. The main effect of Condition was significant because there were differences in $\log(\text{RT})$ overall according to Condition, with Distracter real

words showing the lowest mean log(RT) (Distracter real words: 3.015, CodaS real words: 3.049, Aspirated real words: 3.082, Aspirated non-words: 3.082, Distracter non-words: 3.093, CodaS non-words: 3.095). Finally, the main effect of Group was significant because there were overall differences in log(RT) according to Group. The group with the lowest (i.e., fastest) logRT was the SEVILLE group ($M = 3.02$, $SD = .084$), followed by the Non-Aspirating group ($M = 3.04$, $SD = .087$), At-home group ($M = 3.057$, $SD = .092$), Study Abroad 2 group ($M = 3.065$, $SD = .092$) and lastly the Study Abroad 1 group ($M = 3.066$, $SD = .098$). The interaction between Group and Condition will now be explored further for the NS data and Time 1 L2 data.

Tables 62 and 63 present the means and standard deviations of each Condition according to Group at Time 1 in both milliseconds and log(RT) respectively. Next, post-hoc pairwise comparisons using the Sidak correction were run in order to compare the significance of differences between Condition according to Group and for each Condition within each Group.

Table 62. Mean response time for each Condition according to Group in milliseconds (Time 1). (Parentheses denote standard deviation)

Condition	SEVILLE	NON-ASP	SA1	SA2	AH	Total
CodaS real words	1082.11 (192.08)	1112.13 (203.55)	1187.20 (236.16)	1192.22 (224.66)	1184.59 (221.82)	1149.39 (220.71)
CodaS non-words	1133.49 (197.23)	1234.17 (206.89)	1327.13 (254.22)	1322.44 (235.00)	1303.06 (250.59)	1242.29 (206.89)
Asp real words	1098.37 (198.30)	1199.31 (227.96)	1313.82 (228.96)	1297.72 (252.80)	1282.78 (220.98)	1200.35 (239.92)
Asp non-words	1144.15 (196.78)	1197.63 (222.23)	1294.76 (271.28)	1271.72 (235.08)	1221.39 (228.81)	1219.57 (236.06)
Distracter real words	971.68 (197.57)	1004.46 (212.42)	1049.87 (233.34)	1051.27 (214.24)	1052.62 (221.30)	1023.31 (217.55)
Distracter non-words	1103.58 (188.17)	1155.41 (206.02)	1268.72 (252.39)	1256.31 (238.45)	1227.12 (240.33)	1190.60 (232.92)

Table 63. Mean log(RT) and standard deviations for each Condition according to Group (Time 1) (parentheses denote estimated marginal means controlled for syllables)

Condition	SEVILLE	NON-ASP	SA1	SA2	AH	Total
CodaS real words	$M = 3.028$ $SD = .077$ (3.021)	$M = 3.039$ $SD = .077$ (3.034)	$M = 3.066$ $SD = .087$ (3.065)	$M = 3.069$ $SD = .081$ (3.069)	$M = 3.066$ $SD = .081$ (3.062)	$M = 3.053$ $SD = .083$ (3.050)
CodaS non-words	$M = 3.048$ $SD = .076$ (3.049)	$M = 3.085$ $SD = .073$ (3.077)	$M = 3.115$ $SD = .085$ (3.122)	$M = 3.114$ $SD = .079$ (3.113)	$M = 3.107$ $SD = .087$ (3.107)	$M = 3.086$ $SD = .085$ (3.094)
Asp real words	$M = 3.034$ $SD = .079$ (3.032)	$M = 3.071$ $SD = .083$ (3.066)	$M = 3.112$ $SD = .077$ (3.108)	$M = 3.105$ $SD = .087$ (3.107)	$M = 3.102$ $SD = .076$ (3.099)	$M = 3.071$ $SD = .087$ (3.082)
Asp non-words	$M = 3.052$ $SD = .076$ (3.050)	$M = 3.071$ $SD = .080$ (3.071)	$M = 3.102$ $SD = .093$ (3.109)	$M = 3.097$ $SD = .082$ (3.099)	$M = 3.079$ $SD = .081$ (3.080)	$M = 3.078$ $SD = .084$ (3.082)
Distracter real words	$M = 2.979$ $SD = .086$ (2.991)	$M = 2.993$ $SD = .086$ (3.007)	$M = 3.011$ $SD = .093$ (3.030)	$M = 3.013$ $SD = .085$ (3.032)	$M = 3.013$ $SD = .087$ (3.031)	$M = 3.001$ $SD = .089$ (3.018)
Distracter non-words	$M = 3.037$ $SD = .074$ (3.056)	$M = 3.056$ $SD = .076$ (3.072)	$M = 3.095$ $SD = .087$ (3.120)	$M = 3.091$ $SD = .083$ (3.098)	$M = 3.013$ $SD = .087$ (3.108)	$M = 3.068$ $SD = .084$ (3.091)

CodaS conditions. First, within the CodaS real word condition, the SEVILLE group's mean RT ($M = 1082.11$, $SD = 192.08$) was significantly faster than the Study Abroad 1 ($M = 1187.20$, $SD = 236.16$, $p < .01$), Study Abroad 2 ($M = 1192.22$, $SD = 224.82$, $p < .01$) and At-home groups ($M = 1184.59$, $SD = 221.82$, $p < .05$), but not the Non-Aspirating group ($M = 1112.13$, $SD = 203.55$, $p = .997$). On the other hand, the Non-Aspirating group was not significantly different from the three L2 groups ($p = .357$ to $.682$), and thus was between the SEVILLE and L2 groups. Among the L2 groups in the CodaS real word condition, there were no significant differences ($p = .357$ to 1.00). The same holds true in the CodaS non-word condition. The SEVILLE group ($M = 1133.49$, $SD = 197.23$) was significantly faster than the Study Abroad 1 ($M = 1327.13$, $SD = 254.22$, $p < .001$), Study Abroad 2 ($M = 1322.44$, $SD = 235.00$, $p < .001$) and At-home groups ($M = 1303.06$, $SD = 250.59$, $p < .001$), but not the Non-Aspirating group ($M = 1234.17$, $SD = 206.89$, $p = .549$). The Non-Aspirating group was not significantly different from the L2 groups ($p = .066$ to $.457$) and the L2 groups were not significantly different from one other ($p = .958$ to 1.00).

Asp conditions. In the aspirated condition, beginning with the real words, the SEVILLE group ($M = 1098.37$, $SD = 198.30$) was again significantly faster at responding than all three L2 groups ($p < .001$ for all pairwise comparisons). The Non-Aspirating group ($M = 1199.31$, $SD = 227.96$) was not significantly different from the SEVILLE group ($p = .324$), nor from any of the L2 groups ($p = .176$ to $.482$). Finally, the Study Abroad 1 ($M = 1313.82$, $SD = 228.96$), Study Abroad 2 ($M = 1297.72$, $SD = 252.80$) and At-home ($M = 1200.35$, $SD = 239.92$) groups were not significantly different from one another ($p = .308$ to 1.00). In the Asp non-word condition, the SEVILLE group ($M = 1144.15$, $SD = 196.78$) was significantly faster than the Study Abroad 1 ($M = 1294.76$, $SD = 271.28$, $p < .001$) and Study Abroad 2 ($M = 1271.72$, $SD = 235.08$, $p < .001$) groups, and was approaching significance in being faster than the At-home group ($M = 1221.39$, $SD = 228.81$, $p = .081$). Again, the SEVILLE group was not significantly different from the Non-Aspirating group ($M = 1197.63$, $SD = 222.23$, $p = .899$) and the Non-Aspirating group was not significantly different from the three L2 groups ($p = .221$ to $.999$). Finally, the three L2 groups were not significantly different from one another ($p = .308$ to 1.00).

Distracter conditions. In the first of the final two conditions, the distracter real word condition, a very similar pattern holds. The SEVILLE group ($M = 971.68$, $SD = 197.57$) was significantly faster at responding than all three L2 groups ($p < .05$ for all pairwise comparisons) and was not significantly different from the Non-Aspirating group ($M = 1004.46$, $SD = 212.42$, $p = .963$). The Non-Aspirating group, again, was not significantly different from all other groups ($p = .917$ to $.963$). The Study Abroad 1 ($M = 1049.87$, $SD = 233.34$), Study Abroad 2 ($M = 1051.27$, $SD = 214.30$) and At-home ($M = 1052.62$, $SD = 221.30$) groups were not significantly different ($p = 1.00$ for all pairwise comparisons). The results for the Distracter non-word condition are very similar. The SEVILLE group ($M = 1103.85$, $SD = 188.17$) was significantly faster than all three L2 groups ($p < .001$ for all pairwise comparisons) and not different from the Non-Aspirating group ($M = 1155.41$, $SD = 206.02$, $p = .966$). The Non-Aspirating group, this time, was not significantly different from the At-home group ($M = 1227.12$, $SD = 240.33$, $p = .323$), was approaching

significance in being faster than the Study Abroad 2 group ($M = 1256.31$, $SD = 238.45$, $p = .099$), and was significantly faster than the Study Abroad 1 group ($M = 1268.72$, $SD = 252.39$, $p < .05$).

Summary. The results for Condition according to Group for the NSs and Time 1 L2 learner data showed that the SEVILLE group was significantly faster at responding to all conditions than all of the L2 listener groups, but was not significantly faster than the Non-Aspirating group in any condition. The Non-Aspirating group was in between the SEVILLE group and L2 groups, as it was not significantly different from any of the groups, except for being significantly faster than the Study Abroad 1 group in the non-word distracter condition. Finally, the L2 groups were not significantly different from one another in any of the conditions at Time 1.

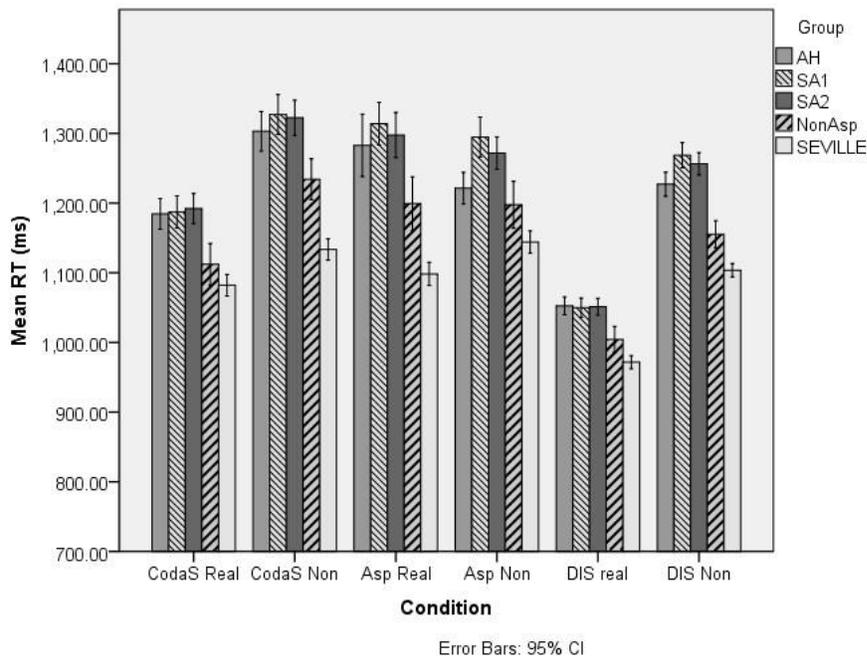


Figure 16. Summary of RTs for Condition according to Group (Time 1)

Results for each Group according to Condition (Time 1)

Now, the results for each group according to Condition are explored, beginning with the SEVILLE group, followed by the Non-Aspirating, Study Abroad 1, Study Abroad 2 and At-home groups (Time 1 for

the L2 groups). Recall that the two NS groups only completed the task once, and thus are reported as a baseline for comparison with the L2 groups. Post-hoc comparisons of Group according to Coda Condition were run using the Sidak correction.

SEVILLE group. First, looking at the CodaS real word condition, the SEVILLE group's mean response times ($M = 1082.11$, $SD = 192.08$) were significantly faster than in the CodaS non-word, Asp non-word, and Distracter non-word conditions ($p < .001$). However, the CodaS RTs were not different from the Asp real word condition ($M = 1098.37$, $SD = 198.30$, $p = .848$) and the Distracter real word condition ($M = 971.68$, $SD = 197.57$), compared to which the CodaS condition resulted in slower RTs ($p < .001$). The mean RT for the Asp real word condition was also not significantly different from that of the Distracter non-word condition ($M = 1103.58$, $SD = 188.17$, $p = 1.00$) and the CodaS non-word condition ($M = 1133.49$, $SD = 172.23$, $p = .099$), though the difference was approaching significance. RTs in the Asp real word condition, though, were significantly slower than those of the Distracter real word condition ($p < .001$), which was the condition that resulted in faster RTs than all other conditions ($p < .001$ for all pairwise comparisons). In nearly all cases, RTs for non-word conditions were significantly slower than for their real word counterparts. This was true of the CodaS real word and non-word conditions ($p < .001$) and Distracter conditions ($p < .001$), but not for the Asp real word and non-word conditions ($p = .074$), though it was approaching significance.

Non-Aspirating group. The Non-Aspirating group's patterns among the Conditions were similar to those of the SEVILLE group. As for the SEVILLE group, the lowest mean RT was found in the Distracter real word condition ($M = 1004.46$, $SD = 212.42$), which was significantly faster than all other conditions (CodaS real word: $p < .01$; all others $p < .001$). The CodaS real word condition ($M = 1112.13$, $SD = 203.55$), apart from slower RTs than the Distracter real word condition, resulted in RTs that were significantly faster than for all other conditions (Asp real word: $p < .01$; CodaS non-word/Asp non-word/Distracter non-word: $p < .001$). As for the SEVILLE group, the CodaS and Distracter real words

resulted in faster RTs on average compared to their non-word counterparts. The mean RT for the Asp real word condition ($M = 1199.31$, $SD = 227.96$), on the other hand, was not significantly different from those of the CodaS non-word ($M = 1234.17$, $SD = 206.89$, $p = .971$), Asp non-word ($M = 1197.63$, $SD = 222.23$, $p = 1.00$) and Distracter non-word ($M = 1155.41$, $SD = 206.02$, $p = 1.00$) conditions. With the exception of the Asp conditions, real words were accessed from the mental lexicon faster than non-words.

Study Abroad 1 group (Time 1). At Time 1, the Study Abroad 1 group's mean RTs were not significantly different among the CodaS non-word ($M = 1327.13$, $SD = 254.22$), Asp real word ($M = 1313.82$, $SD = 228.96$), Asp non-word ($M = 1294.76$, $SD = 271.28$), and Distracter non-word ($M = 1268.72$, $SD = 252.39$) conditions ($p = .314$ to 1.00). The Distracter real word condition ($M = 1049.87$, $SD = 233.34$), as it did for the NS groups, resulted in the fastest RTs of all groups ($p < .001$). The CodaS real word condition ($M = 1187.20$, $SD = 236.16$) was second and RTs were significantly faster than in all other conditions except for the Distracter real word condition (Asp real word: $p < .05$; all others: $p < .001$).

Study Abroad 2 group (Time 1). The Study Abroad 2 group followed a nearly identical pattern to the Study Abroad 1 group at Time 1. As for all other groups thus far, the Distracter real word ($M = 1051.27$, $SD = 214.24$) condition exhibited the fastest response times of all conditions by a significant margin ($p < .001$ for all pairwise comparisons). Like the Study Abroad 1 group, the CodaS non-word ($M = 1322.44$, $SD = 235.00$), Asp real word ($M = 1297.72$, $SD = 252.80$), Asp non-word ($M = 1271.72$, $SD = 235.08$) and Distracter non-word ($M = 1256.31$, $SD = 238.45$) conditions were not significantly different from one another ($p = .249$ to 1.00). The CodaS real word condition ($M = 1192.22$, $SD = 224.66$) was significantly slower than the Distracter real word condition ($p < .001$), but significantly faster than all other conditions ($p < .001$ for all pairwise comparisons).

At-home group (Time 1). Finally, the At-home group was also nearly identical to both study abroad groups at Time 1. First, like all other groups, the mean RT in the Distracter real word condition

($M = 1052.62$, $SD = 221.30$) was significantly faster than for all other conditions ($p < .001$). Also, the CodaS real word condition exhibited faster RTs than the CodaS non-word ($M = 1303.06$, $SD = 250.59$, $p < .001$), Asp real word ($M = 1282.78$, $SD = 220.98$, $p < .01$), Asp non-word ($M = 1221.39$, $SD = 228.81$, $p < .05$) and Distracter non-word ($M = 1227.12$, $SD = 240.33$, $p < .001$) conditions. Finally, the Asp real word, CodaS non-word, and Distracter non-word conditions were again not significantly different from one another ($p = .600$ to 1.00), but the Asp non-word condition did exhibit significantly faster RTs than the Distracter non-word condition ($p < .01$).

Summary. The patterns for each group according to Condition are similar in some ways and different in others. Each group was the fastest at responding to the Distracter real words and the second fastest at responding to the CodaS real words. The SEVILLE group showed a consistent pattern across real words and non-words in that they responded to the real words significantly faster than the non-words (except that the comparison of Asp real words and Asp non-words was approaching significance at $p = .074$). Importantly, the SEVILLE group also responded to Asp real words with the same speed as CodaS real words, and was the only group to do so. The two study abroad groups, At-home group and Non-Aspirating group all patterned like the SEVILLE group for the CodaS real word and non-words and the Distracter real word and non-words. All the while, the responses to Asp real words were always significantly slower than the CodaS real words and were not different from the Asp non-words.

Results for each Condition according to Group (Time 2)

In order to compare the groups' RTs according to Condition at Time 2, the Time 1 data was excluded and a linear mixed effects model was run in which Subject and Item were random effects and Group, Condition and Syllables were fixed effects. It is important to remember that the NS data is the same as at Time 1, as they were only measured once. The model showed significant main effects for Syllables ($F(1, 351.20) = 121.393$, $p < .001$), Condition ($F(5, 365.11) = 67.715$, $p < .001$) and Group ($F(4,$

112.52) = 3.556, $p < .01$), as well as a significant interaction effect of Condition*Group ($F(20, 16540.12) = 5.484$, $p < .001$). The main effect of Syllables was significant because the log(RT) for two-syllable words was significantly lower (i.e., faster) than the log(RT) for three-syllable words. The main effect of Condition was significant because there were significant differences in log(RT) according to Condition. The order of mean log(RT) for each condition, from lowest to highest, is Distracter real words (3.007), CodaS real words (3.043), Aspirated real words (3.068), Aspirated non-words (3.070), Distracter non-words (3.078), and CodaS non-words (3.079). Finally, the main effect of Group was significant because, while most of the groups were not significantly different ($p = .146$ to 1.00) in terms of overall log(RT) regardless of condition, the SEVILLE group's mean log(RT) was significantly faster than that of Study Abroad 2 ($p < .01$). The significant interaction between Group and Condition will now be explored.

Tables 64 and 65 present the means and standard deviations of each Condition according to Group at Time 2, first for the millisecond data and then for the log(RT) data. Next, Post-hoc pairwise comparisons using the Sidak correction were run in order to compare the significance of differences between Condition according to Group and for each Condition within each Group.

Table 64. Mean response time (ms) for each Condition according to Group (Time 2). (Parentheses denote standard deviation)

Condition	SEVILLE	NON-ASP	SA1	SA2	AH	Total
CodaS real words	1082.11 (192.08)	1112.13 (203.55)	1135.91 (244.93)	1181.41 (233.32)	1167.83 (229.46)	1133.69 (224.19)
CodaS non-words	1133.49 (197.23)	1234.17 (206.89)	1228.64 (256.31)	1265.73 (246.42)	1224.68 (241.73)	1206.08 (234.96)
Asp real words	1098.37 (198.30)	1199.31 (227.96)	1182.49 (247.95)	1247.55 (231.21)	1284.26 (220.90)	1170.79 (230.66)
Asp non-words	1144.15 (196.78)	1197.63 (222.23)	1216.41 (253.02)	1258.60 (242.86)	1167.39 (232.35)	1191.69 (232.08)
Distracter real words	971.68 (197.57)	1004.46 (212.42)	1003.33 (232.41)	1036.22 (216.11)	1026.22 (218.37)	1005.66 (215.69)
Distracter non-words	1103.58 (188.17)	1155.41 (206.02)	1177.12 (251.84)	1216.85 (241.59)	1167.20 (242.28)	1158.35 (228.20)

Table 65. Mean log(RT) and standard deviations for each Condition according to Group (Time 2). (Parentheses denote estimated marginal means controlled for syllables)

Condition	SEVILLE	NON-ASP	SA1	SA2	AH	Total
CodaS real words	$M = 3.028$ $SD = .077$ (3.021)	$M = 3.039$ $SD = .077$ (3.034)	$M = 3.046$ $SD = .092$ (3.045)	$M = 3.064$ $SD = .085$ (3.065)	$M = 3.059$ $SD = .085$ (3.058)	$M = 3.046$ $SD = .085$ (3.045)
CodaS non-words	$M = 3.048$ $SD = .076$ (3.049)	$M = 3.085$ $SD = .073$ (3.077)	$M = 3.080$ $SD = .091$ (3.080)	$M = 3.094$ $SD = .084$ (3.096)	$M = 3.080$ $SD = .085$ (3.081)	$M = 3.073$ $SD = .084$ (3.077)
Asp real words	$M = 3.034$ $SD = .079$ (3.032)	$M = 3.071$ $SD = .083$ (3.066)	$M = 3.063$ $SD = .091$ (3.066)	$M = 3.089$ $SD = .080$ (3.088)	$M = 3.102$ $SD = .079$ (3.081)	$M = 3.060$ $SD = .086$ (3.066)
Asp non-words	$M = 3.052$ $SD = .076$ (3.050)	$M = 3.071$ $SD = .080$ (3.071)	$M = 3.076$ $SD = .091$ (3.074)	$M = 3.092$ $SD = .084$ (3.093)	$M = 3.059$ $SD = .085$ (3.058)	$M = 3.068$ $SD = .085$ (3.069)
Distracter real words	$M = 2.979$ $SD = .086$ (2.991)	$M = 2.993$ $SD = .086$ (3.007)	$M = 2.991$ $SD = .095$ (3.009)	$M = 3.007$ $SD = .087$ (3.025)	$M = 3.002$ $SD = .087$ (3.023)	$M = 2.993$ $SD = .089$ (3.011)
Distracter non-words	$M = 3.037$ $SD = .074$ (3.056)	$M = 3.056$ $SD = .076$ (3.072)	$M = 3.061$ $SD = .092$ (3.082)	$M = 3.077$ $SD = .085$ (3.098)	$M = 3.058$ $SD = .088$ (3.083)	$M = 3.056$ $SD = .084$ (3.078)

CodaS conditions. For the Time 2 L2 learner data compared to the NS data, within the CodaS real word condition, the SEVILLE group's mean RT ($M = 1082.11$, $SD = 192.08$) was significantly faster than the Study Abroad 2 group ($M = 1181.41$, $SD = 233.32$, $p < .05$), but was not faster than other groups ($p = .108$ to $.998$). This is a difference compared to the Time 1 data, when the SEVILLE group was not significantly faster only than the Non-Aspirating group in the CodaS real word condition. However, the Non-Aspirating group was still not significantly different from the three L2 groups or the SEVILLE group ($p = .692$ to 1.00) at Time 2 for the L2 groups. Among the L2 groups in the CodaS real word condition, there were still no significant differences at Time 2 ($p = .804$ to $.999$). In the CodaS non-word condition, The SEVILLE group ($M = 1133.49$, $SD = 197.23$) was significantly faster only than the Study Abroad 2 group ($M = 1265.73$, $SD = 246.42$, $p < .01$) while the Study Abroad 1 ($M = 1228.64$, $SD = 256.31$) and At-home groups ($M = 1224.68$, $SD = 241.73$) were no longer significantly slower than the SEVILLE group, but were approaching significance ($p = .089$ and $.075$ respectively). The Non-Aspirating group was still not

significantly different from the L2 groups ($p = .931$ to 1.00) and the L2 groups were not significantly different from one another ($p = .960$ to $.962$).

Asp conditions. In the Asp real word condition, there was a difference among the L2 groups compared to the SEVILLE group ($M = 1098.37$, $SD = 198.30$). At Time 2 for the L2 learners, the SEVILLE group's RTs were only significantly faster than the Study Abroad 2 ($M = 1247.55$, $SD = 231.21$, $p < .001$) and At-home ($M = 1284.26$, $SD = 220.90$, $p < .01$) groups in this condition, while the Study Abroad 1 group ($M = 1182.49$, $SD = 247.95$) was only approaching significance due to having speeded response times to a level that was nearly on par statistically-speaking with the SEVILLE group at Time 2 ($p = .079$). The Non-Aspirating group was still not significantly different from any of the other groups in this condition ($p = .401$ to 1.00). Finally, the Study Abroad 1, Study Abroad 2 and At-home groups were not significantly different from one another at Time 2 ($p = .803$ to 1.00). This is a case in which the difference between Study Abroad 1 ($M = 1182.49$, $SD = 247.95$) and the At-home group ($M = 1284.26$, $SD = 220.90$, $p < .01$) in the Asp real word condition appears that it should be significant, but is not when the number of syllables is controlled for by the model. It is clear, too, that the At-home group's 95% confidence interval is much larger than that of the study abroad groups, which also contributes to an insignificant difference (see error bars of Figure 17). In the non-word Asp condition, the SEVILLE group ($M = 1144.15$, $SD = 196.78$) was significantly faster only than the Study Abroad 2 group ($M = 1258.60$, $SD = 242.86$, $p < .01$), but was not significantly faster than the other groups ($p = .482$ to 1.00). The Non-Aspirating group was not significantly different from the three L2 groups ($p = .822$ to 1.00). Finally, the Study Abroad 1 and Study Abroad 2 groups were not significantly different from one another ($p = .783$), while the Study Abroad 2 and At-home groups were on the verge of being significantly different ($p = .051$). Specifically, the At-home group was faster than the Study Abroad 2 group in the Asp non-word condition.

Distracter conditions. In the first of the final two conditions, the Distracter real word condition, a different pattern obtains at Time 2 for the L2 groups than at Time 1. The SEVILLE group ($M = 971.68$, SD

= 197.57) was no longer significantly faster at responding to Distracter real word items than any other group ($p = .200$ to $.979$). The Non-Aspirating group, again, was not significantly different from all other groups ($p = .979$ to 1.00). The Study Abroad 1 ($M = 1003.33$, $SD = 216.11$), Study Abroad 2 ($M = 1036.22$, $SD = 216.11$) and At-home ($M = 1026.22$, $SD = 218.37$) groups were not significantly different ($p = .954$ to $.986$). The results for the Distracter non-word condition show that the SEVILLE group ($M = 1103.85$, $SD = 188.17$) was significantly faster only than the Study Abroad 2 group ($M = 1216.85$, $SD = 241.59$) and not different from the Non-Aspirating group ($M = 1155.41$, $SD = 206.02$, $p = .986$), Study Abroad 1 group ($M = 1177.12$, $SD = 251.84$, $p = .320$), or At-home group ($M = 1167.20$, $SD = 242.28$, $p = .298$). The Non-Aspirating group was not significantly different from any other group ($p = .746$ to 1.00).

Summary. The analysis of Condition according to Group at Time 2 revealed different patterns than at Time 1. Most notably, the SEVILLE group was no longer significantly faster in all conditions than every other group. Of the L2 groups, the Study Abroad 1 group speeded RTs the most at Time 2 so that it was not significantly different from the SEVILLE group in any condition, though there were a few conditions in which the Study Abroad 1 group's slower means compared to the SEVILLE group approached significance. The Study Abroad 2 group, on the other hand, was still significantly slower than the SEVILLE group at Time 2 for all but the distracter conditions. Additionally, the Study Abroad 2 group was very nearly slower than the At-home group in the Asp non-word condition, but the three L2 groups were not significantly different in any other condition, including the Asp real word condition.

Continuing with the Asp real word condition, the SEVILLE group was still significantly faster than the Study Abroad 2 and At-home groups, but was almost not significantly different from the Study Abroad 1 group. Like at Time 1, the Non-Aspirating group was not significantly different from any other group in any condition and was intermediate between the L2 and SEVILLE groups. Figure 17 displays a summary of the mean RTs in all Conditions according to Group at Time 2. Next, the pairwise comparisons for the RTs in each condition within each group will be presented.

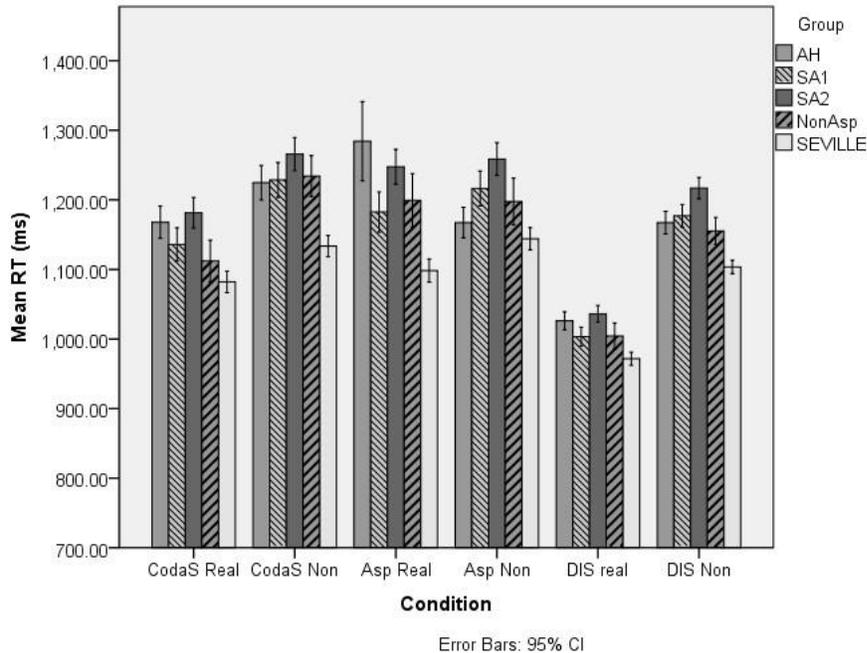


Figure 17. Summary of RTs for Condition according to Group (Time 2)

Results for each Group according to Condition (Time 2)

Study Abroad 1 group (Time 2). In presenting the Time 2 results for Group according to Condition, the NS groups will not be repeated except for in the summary tables and figures as a comparison with the L2 groups. At Time 2, the Study Abroad 1 group was still fastest at responding to Distracter real words ($M = 1003.33, SD = 232.41$) than any other condition ($p < .001$). The CodaS real word condition ($M = 1135.91, SD = 244.93$), too, still had significantly faster RTs than all other conditions (Asp real word: $p < .05$; all others: $p < .001$) except the Distracter real word condition. Finally, the Asp real word ($M = 1182.49, SD = 247.95$), CodaS non-word ($M = 1228.64, SD = 256.31$), Asp non-word ($M = 1216.41, SD = 253.02$) and Distracter non-word ($M = 1177.12, SD = 251.84$) conditions were not significantly different from one another ($p = .314$ to 1.00). The pattern across conditions at Time 2 is more like that of the SEVILLE group than it was at Time 1 in that there was a trend for non-words to be accessed more slowly than real words.

Study Abroad 2 group (Time 2). The Study Abroad 2 group followed an identical pattern at Time 2 compared to Time 1 in terms of the significance of differences between the conditions. The Distracter real word ($M = 1036.22$, $SD = 216.11$) condition again exhibited the fastest response times of all conditions by a significant margin ($p < .001$ for all pairwise comparisons). Like at Time 1, RTs in the CodaS non-word ($M = 1265.73$, $SD = 246.42$), Asp real word ($M = 1247.55$, $SD = 231.21$), Asp non-word ($M = 1258.60$, $SD = 242.86$) and Distracter non-word ($M = 1216.85$, $SD = 241.59$) conditions were not significantly different from one another ($p = .858$ to 1.00). Finally, though RTs in the CodaS real word condition ($M = 1192.22$, $SD = 224.66$) were significantly slower than those in the Distracter real word condition ($p < .001$), they were significantly faster than in all other conditions (Asp real word: $p < .05$; all others: $p < .001$). Like the Study Abroad 1 group, the Study Abroad 2 group's pattern of RTs for real words compared to non-words was beginning to look more like the SEVILLE group at Time 2, but the Asp real word condition was still nearly equal to the Asp non-word condition.

At-home group (Time 2). Finally, the At-home group was similar at Time 2 compared to Time 1, but with a few differences. First, like at Time 1, the mean RT in the Distracter real word condition ($M = 1026.22$, $SD = 218.37$) was significantly faster than for all other conditions ($p < .001$). Also, the CodaS real word condition ($M = 1167.83$, $SD = 229.46$), exhibited faster RTs than the CodaS non-word ($M = 1224.68$, $SD = 241.73$, $p < .01$) and Distracter non-word ($M = 1167.20$, $SD = 242.28$, $p < .001$) conditions, but not the Asp non-word ($M = 1167.39$, $SD = 232.35$, $p = 1.00$) and Asp real word conditions ($M = 1282.78$, $SD = 220.98$, $p = .06$), though the latter was approaching significance. Finally, the Asp real word, CodaS non-word, and Distracter non-word conditions were again not significantly different from one another ($p = 1.00$), but the Asp non-word condition did exhibit significantly faster RTs than the Distracter non-word condition ($p < .01$).

Summary. At Time 2, the L2 groups followed similar trends to Time 1 overall. For all groups, the Distracter real word condition had the fastest RTs, followed by the CodaS real word condition. For the

At-home group, the CodaS condition exhibited RTs that were faster than the CodaS non-word and Distracter non-word conditions, but not the Asp non-word and Asp real word conditions. For all three groups, the Asp real word, CodaS non-word, and Distracter non-word conditions showed RTs that were not significantly different. For the Study Abroad 1 and Study Abroad 2 groups, the Asp non-word condition was also not significantly different from Asp real word, CodaS non-word, and Distracter non-word conditions. Interestingly, although the Study Abroad 1 and Study Abroad 2 groups accessed Asp real words more slowly than they rejected Asp non-words at Time 1, they were beginning to access the Asp real words more quickly than they rejected Asp non-words at Time 2. This was particularly true of the Study Abroad 1 group, which was nearly not different from the SEVILLE group, statistically-speaking, at Time 2 in the Asp real word condition. The SEVILLE group accessed real words faster than it rejected non-words in all conditions and the study abroad groups were beginning to follow this same pattern at Time 2. Figure 18 presents a summary of the result of each Group according to Condition at Time 2, while the At-home group did not.

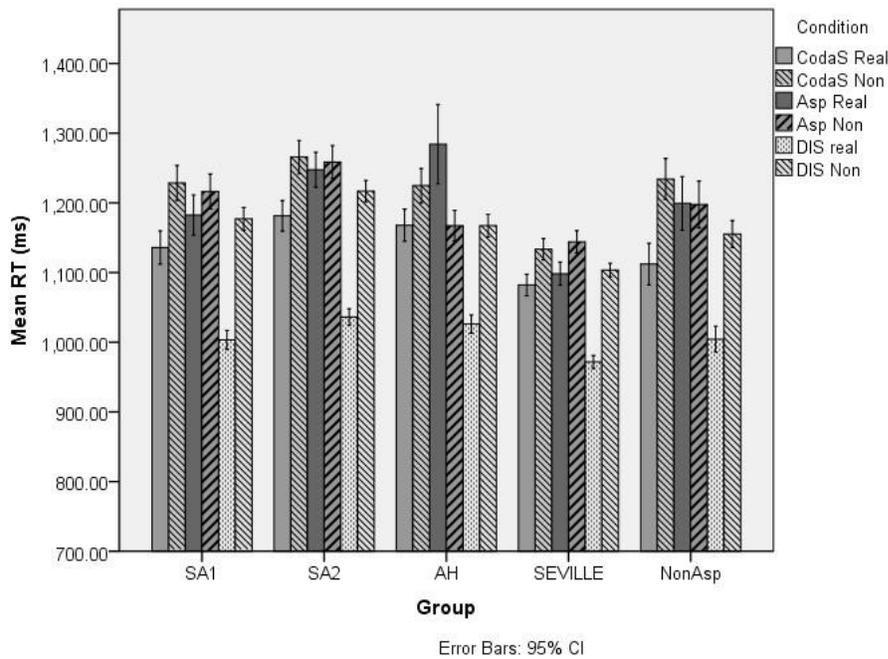


Figure 18. RTs for each Group according to Condition at Time 2 (N=116)

Analysis of the Effect of Time (i.e., Exposure) on the L2 Learners' RTs

Now that the data from Time 1 and Time 2 have been presented separately, revealing the overall patterns of each group in each coda condition at each data collection time (i.e., twice for the L2 groups and once for the NS groups), the effect of the variable Time on the L2 learners' response times will be examined in order to test whether exposure to Andalusian /s/-aspiration over time played a significant role in the response times in the different conditions for the three L2 groups. To accomplish this, a linear mixed effects model was run to evaluate the change in accuracy by coda condition according to time and between groups. Subject and Item were included as random effects and the fixed effects included were Time (Time 1, Time 2), Group (Study Abroad 1, Study Abroad 2, At-home), Condition (CodaS real words, CodaS non-words, Asp real words, Asp non-words, Distracter real words, Distracter non-words), and Syllables along with their interactions. The model showed significant main effects for Syllables ($F(1, 262.993) = 106.646, p < .001$), Time ($F(1, 19783.36) = 271.738, p < .001$), Condition ($F(5, 267.69) = 51.66, p < .001$), but no significant main effect of Group ($F(2, 70.759) = .385, p = .682$). The main effect of Syllables was significant because the mean $\log(\text{RT})$ for two-syllable words was significantly lower (i.e., faster) than for three-syllable words. The main effect for Condition was significant because of differences in $\log(\text{RT})$ among the conditions. For example, $\log(\text{RT})$ in the Distracter real word and CodaS real word conditions was significantly lower than all other conditions ($p < .01$ to $.001$). However, the overall mean $\log(\text{RT})$ in the Aspirated real word condition, Aspirated non-word condition, and Distracter non-word condition was not significantly different. The main effect of Time was significant because the mean $\log(\text{RT})$ of all groups and conditions combined was significantly lower (i.e., faster) at Time 2 than at Time 1. However, the mean $\log(\text{RT})$ for the three L2 groups overall (i.e., regardless of Condition and Time) was not significantly different, which is why there was no main effect of Group.

There were also significant interaction effects for Condition*Group ($F(10, 19729.23) = 8.048, p < .001$), Condition*Time ($F(5, 19761.09) = 100.818, p < .001$), and Group*Time ($F(2, 19744.90) = 32.748, p < .001$), but no significant interaction effect for Group*Condition*Time ($F(10, 19712.31) = .857, p = .573$). The interaction effect of Group*Time showed that, while there were no significant differences in log(RT) among the groups (i.e., regardless of Condition) at Time 1 and Time 2, there were significant differences in log(RT) between Time 1 and Time 2 within each group overall. The interaction effect of Condition*Group showed that, while there were no significant differences of log(RT) among the three groups within each condition (i.e., regardless of Time), there were significant differences of log(RT) among the six conditions within each group. Finally, the interaction effect of Condition*Time was significant because there was significant change over time of log(RT) within each condition (i.e., regardless of Group), and that there were significant differences of log(RT) among the six conditions at both timepoints.

The three-way interaction of Group*Condition*Time was not statistically significant because there were no significant differences of log(RT) among the three L2 groups at Time 1 or Time 2 except for the Aspirated non-word condition at Time 2. But, while the interaction overall was not significant, it is important to look at what occurred over time in the different conditions, particularly the Asp real word condition, because an interesting finding emerges. In order to do this, each group's results for both Time 1 and Time 2 for each condition are presented, followed by a summary. Table 66 summarizes the RTs of each L2 group in each condition according to Time.

Table 66. Comparison of the L2 groups' RTs (ms) in each Condition according to Time (parentheses denote standard deviation)

Condition*Time	SA1	SA2	AH
CodaS real words Time 1	1187.20 (236.16)	1192.22 (224.66)	1184.59 (221.82)
CodaS real words Time 2	1135.91 (244.93)	1181.41 (233.32)	1167.83 (229.46)
CodaS non- words Time 1	1327.13 (254.22)	1322.44 (235.00)	1303.06 (250.59)
CodaS non-words Time 2	1228.64 (256.31)	1265.73 (246.42)	1224.68 (241.73)
Aspirated real words Time 1	1313.82 (228.96)	1297.72 (252.80)	1282.78 (220.98)
Aspirated real words Time 2	1182.49 (247.95)	1247.55 (231.21)	1284.26 (220.90)
Aspirated non-words Time 1	1294.76 (271.28)	1271.72 (235.08)	1221.39 (228.81)
Aspirated non-words Time 2	1216.41 (253.02)	1258.60 (242.86)	1167.39 (232.35)
Distracter real words Time 1	1049.87 (233.34)	1051.27 (214.24)	1052.62 (221.30)
Distracter real words Time 2	1003.33 (232.41)	1036.22 (216.11)	1026.22 (218.37)
Distracter non-words Time 1	1268.72 (252.39)	1256.31 (238.45)	1227.12 (240.33)
Distracter non-words Time 2	1177.12 (251.84)	1216.85 (241.59)	1167.20 (242.28)

Table 67. Comparison of the L2 groups' log(RT) in each Condition according to Time (parentheses denote estimated marginal means for log(RT) controlled for syllables)

Condition*Time	SA1	SA2	AH
CodaS real words Time 1	<i>M</i> = 3.066 <i>SD</i> = .087 (3.065)	<i>M</i> = 3.069 <i>SD</i> = .081 (3.069)	<i>M</i> = 3.066 <i>SD</i> = .081 (3.062)
CodaS real words Time 2	<i>M</i> = 3.046 <i>SD</i> = .092 (3.045)	<i>M</i> = 3.064 <i>SD</i> = .085 (3.065)	<i>M</i> = 3.059 <i>SD</i> = .085 (3.058)
CodaS non- words Time 1	<i>M</i> = 3.115 <i>SD</i> = .085 (3.122)	<i>M</i> = 3.114 <i>SD</i> = .079 (3.113)	<i>M</i> = 3.107 <i>SD</i> = .087 (3.107)
CodaS non-words Time 2	<i>M</i> = 3.080 <i>SD</i> = .091 (3.080)	<i>M</i> = 3.094 <i>SD</i> = .084 (3.096)	<i>M</i> = 3.080 <i>SD</i> = .085 (3.081)
Aspirated real words Time 1	<i>M</i> = 3.112 <i>SD</i> = .077 (3.108)	<i>M</i> = 3.105 <i>SD</i> = .087 (3.107)	<i>M</i> = 3.102 <i>SD</i> = .076 (3.099)
Aspirated real words Time 2	<i>M</i> = 3.063 <i>SD</i> = .091 (3.066)	<i>M</i> = 3.089 <i>SD</i> = .080 (3.088)	<i>M</i> = 3.102 <i>SD</i> = .079 (3.081)
Aspirated non-words Time 1	<i>M</i> = 3.102 <i>SD</i> = .093 (3.109)	<i>M</i> = 3.097 <i>SD</i> = .082 (3.099)	<i>M</i> = 3.079 <i>SD</i> = .081 (3.080)
Aspirated non-words Time 2	<i>M</i> = 3.076 <i>SD</i> = .091 (3.074)	<i>M</i> = 3.092 <i>SD</i> = .084 (3.093)	<i>M</i> = 3.059 <i>SD</i> = .085 (3.056)
Distracter real words Time 1	<i>M</i> = 3.011 <i>SD</i> = .093 (3.030)	<i>M</i> = 3.013 <i>SD</i> = .085 (3.032)	<i>M</i> = 3.013 <i>SD</i> = .087 (3.031)
Distracter real words Time 2	<i>M</i> = 2.991 <i>SD</i> = .095 (3.009)	<i>M</i> = 3.007 <i>SD</i> = .087 (3.025)	<i>M</i> = 3.002 <i>SD</i> = .087 (3.023)
Distracter non-words Time 1	<i>M</i> = 3.095 <i>SD</i> = .087 (3.120)	<i>M</i> = 3.091 <i>SD</i> = .083 (3.113)	<i>M</i> = 3.013 <i>SD</i> = .087 (3.108)
Distracter non-words Time 2	<i>M</i> = 3.061 <i>SD</i> = .092 (3.082)	<i>M</i> = 3.077 <i>SD</i> = .085 (3.098)	<i>M</i> = 3.058 <i>SD</i> = .088 (3.083)

Study Abroad 1 group (Time 1 vs. Time 2). The comparison of the Study Abroad 1 group's RTs are Time 1 versus Time 2 show that in each Condition, this group responded significantly faster at Time 2 than at Time 1 ($p < .001$ for all pairwise comparisons). This implies that the Study Abroad 1 group did get

significantly faster at responding to Asp real words over time, but that this was not the only condition that saw a significant decrease in RT over time. Figure 19 displays the RTs by Condition at each Time for the study abroad group.

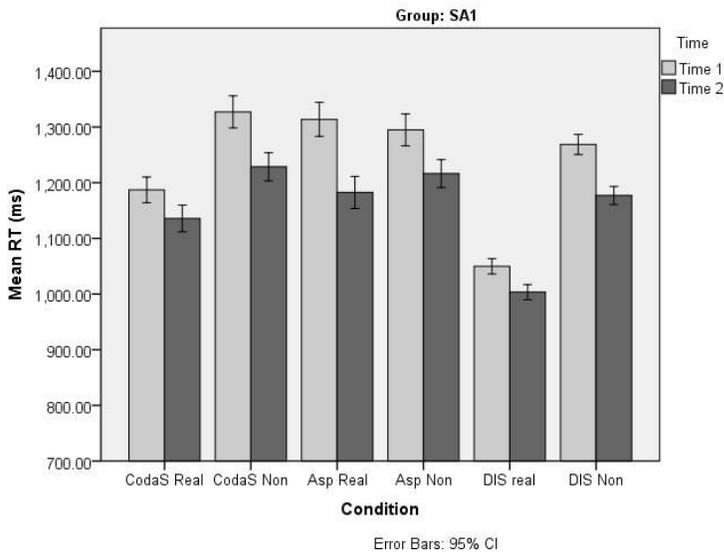


Figure 19. Comparison of the Study Abroad 1 group’s mean RTs in each Condition according to Time.

Study Abroad 2 group (Time 1 vs. Time 2). The Study Abroad 2 group did not follow the same pattern as the Study Abroad 1 group. While the Study Abroad 2 group did evidence significantly faster RTs at Time 2 compared to Time 1 for the CodaS non-word ($p < .01$), Asp real word ($p < .01$), Distracter real word ($p < .05$) and Distracter non-word ($p < .001$) conditions, this was not the case for the CodaS real word ($p = .437$) and Asp non-word ($p = .253$) conditions. Figure 20 displays these results for Time 1 compared to Time 2 for the Study Abroad 2 group.

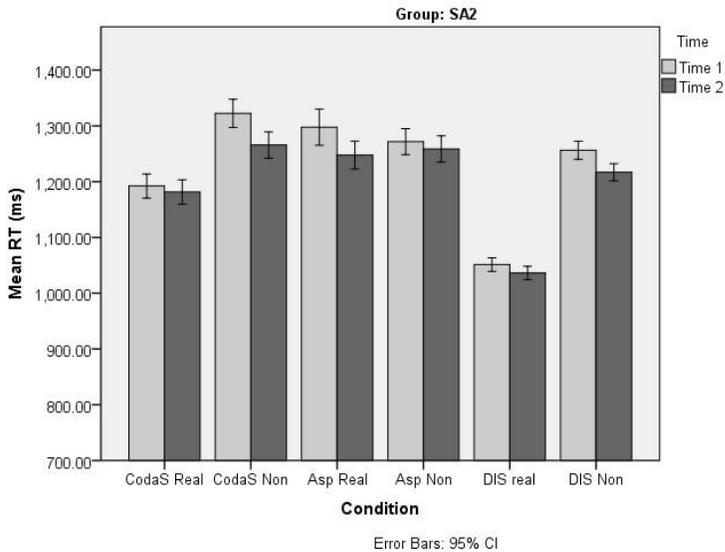


Figure 20. Comparison of the Study Abroad 2 group’s mean RTs in each Condition according to Time.

At-home group (Time 1 vs. Time 2). Finally, the At-home group shows a different pattern than both study abroad groups. The At-home group was significantly faster at Time 2 than Time 1 when responding to CodaS non-words ($p < .001$), Asp non-words ($p < .001$), Distracter real words ($p < .01$) and Distracter non-words ($p < .001$). However, the At-home group was not significantly faster at Time 2 when responding to CodaS real words ($p = .346$) and Asp real words ($p = .138$). Thus, the most important distinction between the At-home group and the two study abroad groups is that the At-home group was the only L2 group that did not evidence significantly faster RTs at Time 2 in the Asp real word condition.

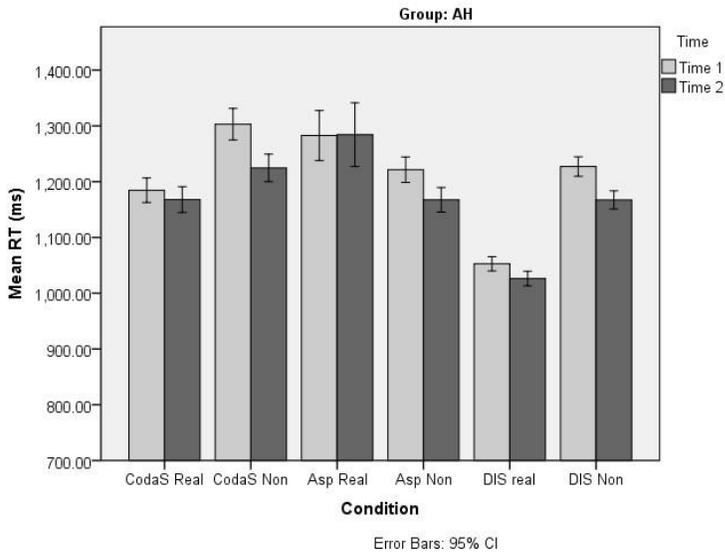


Figure 21. Comparison of the At-home group’s mean RTs in each Condition according to Time.

Summary. A comparison of the three L2 groups’ RTs in each condition at Time 1 and Time 2 revealed that there was variation among the groups as to which conditions showed change in RTs over time. While the Study Abroad 1 group speeded response times at Time 2 to a significant degree in every condition, the Study Abroad 2 and At-home groups varied as to which conditions showed change over time. Both of these groups speeded RTs in four of the six conditions, but the conditions that were significant differed for each group. In other words, there was no consistent set of conditions in which there was change over time for certain groups and not others. Nonetheless, the most interesting finding is that, while the At-home group did speed response times at Time 2 for four of the six conditions, just like the Study Abroad 2 group did, the At-home group was the only L2 group that did not demonstrate significantly faster RTs at Time 2 in the Asp real word condition. The previous analysis of all the groups according to Condition and Time separately showed that the three L2 groups’ RTs in the Asp real word condition at Time 1 and Time 2 were not significantly different. While this is true, a closer analysis has revealed that, while they are not different overall in the Asp real word condition at either Time, only the

study abroad learners, who were exposed to /s/-aspiration during study abroad, responded significantly faster over time to real Spanish words that contained Andalusian /s/-aspiration.

Summary of Chapter 4

In this chapter, the results of the analysis of the two perception tasks were presented, beginning with the accuracy results of the forced-choice identification task and response types and followed by the correlations between accuracy on the identification task and the extralinguistic factors for the study abroad groups. The linear mixed model revealed several differences between the NS groups, study abroad groups and the At-home group for the perception of the seven coda conditions, but the most striking differences were found in the /s/-aspiration condition. The SEVILLE group was clearly the most accurate and the At-home group was the least accurate. The study abroad groups were more accurate than the At-home group at both times and the Non-Aspirating group was intermediate between the L2 groups and the SEVILLE group. It is important to note that, while the study abroad groups were significantly more accurate on both tasks than the At-home group at Time 1, there was still a very clear difference between the groups over time in that only the study abroad groups made significant gains. This will be discussed further in Chapter Five. A few extralinguistic factors were significantly correlated with accuracy among study abroad learners, including some different measures of the amount of Spanish that the learners used overall and also with NSs of Spanish during the semester, showing that the learners who used the most Spanish overall and also spoke the most in Spanish with NSs of Spanish exhibited higher accuracy at Time 2 in the Asp condition. Also, there were different patterns of responses for the aspirated stimuli according to group, as the NS groups showed the least variation and the most association of /s/-aspiration with <s>, while the L2 learners preferred the No Coda response, particularly the At-home group. There was change over time, though, for the study abroad groups, as they decreased association of /s/-aspiration with No Coda over time and increased association with <s>.

The At-home group, though, did not change responses to aspiration over time and still preferred the No Coda response almost categorically at Time 2. The results of the different aspects of the forced-choice identification task are summarized in Table 68 (page 284).

The second part of the chapter presented the results of the analysis of the lexical decision task. The analysis revealed that the two NS groups were very similar except for the Asp real word condition, where the SEVILLE group was significantly faster at responding than the Non-Aspirating group. Also, the Study Abroad 1 and Study Abroad 2 groups' accuracy, which was at chance level at Time 1, increased significantly at Time 2. However, the At-home group's accuracy for aspirated words decreased significantly at Time 2. The At-home group's accuracy was generally lower overall, but the greatest difference between the At-home group and the Study Abroad groups was for the aspirated words. Extralinguistic factors were not as influential in the results of the lexical decision task compared to the forced-choice identification task. But a few factors indicated that learners who read more in Spanish and watched less English media had higher accuracy at Time 2. Also, an interesting finding was that the grammar test score at each data collection time was correlated positively with accuracy for all conditions except the Asp real word condition, showing that this condition was an exception to the rule and served as an equalizer among learners of different grammatical proficiency levels. Finally, response times were analyzed at Time 1 and Time 2. The results showed that the At-home group was the only L2 group that did not get faster at responding to Asp real words at Time 2, while this group did speed response times in several other conditions. The study abroad groups, on the other hand, did speed response times significantly in the Asp real word condition over time. But at Time 2, the three groups were still not significantly different. Of the three L2 groups, though, the Study Abroad 1 group's RTs speeded up the most at Time 2 and this group was similar to the SEVILLE group overall. The results of the different aspects of the lexical decision task are summarized in Table 69 (page 285).

Table 68. Summary of the results of the Identification task (SA = Study Abroad, AH = At-home, TL = target language)

Task	Results
<p>Forced-Choice Identification Accuracy (Time 1)</p>	<ul style="list-style-type: none"> • SEVILLE: 96%-100% for all coda conditions except Asp (74.5%) • NON-ASP: 96%-100% for all coda conditions except Asp (26.1%) • SA1: 95%-100% for all coda conditions except Asp (18.1%) and [f] (81.3%) • SA2: 93%-100% for all coda conditions except Asp (16.7%) and [f] (75.2%) • AH: 90%-100% for all coda conditions except Asp (0.3%) and [f] (79.2%)
<p>Forced-Choice Identification Accuracy (Time 2)</p>	<ul style="list-style-type: none"> • SA1: 96%-100% for all coda conditions except Asp (35.8%) and [f] (87%) • 93%-100% for all coda conditions except Asp (38.2%) and [f] (86%) • 90%-100% for all coda conditions except Asp (0.9%) and [f] (86%)
<p>Forced-Choice Identification Accuracy over time</p>	<ul style="list-style-type: none"> • SA1: Significant positive change only for Asp condition • SA2: Significant positive change for Asp and [f] conditions • AH: Significant positive change only for [f] condition, not Asp
<p>Main Identification Responses to Aspirated Stimuli for Each Group and over Time for L2 Groups</p>	<ul style="list-style-type: none"> • SEVILLE: Asp mapped to <s> 72.5%, [f] 20.5%, No Coda 4.7% • NON-ASP: Asp mapped to <s> 26.1%, [f] 17.9%, No Coda 55.2% • SA1: Asp mapped to <s> 18.1%, No Coda 75.1% @ T1. <s> increased to 35.8% and No Coda reduced to 57.5% @ T2. • SA2: Asp mapped to <s> 16.7% and No Coda 77.9% @ T1. <s> increased to 38.2% and No Coda reduced to 57.9% @ T2. • AH: Asp mapped almost exclusively to No Coda at Time 1 and Time 2 (94%-95%)
<p>Forced-Choice Identification Accuracy for the Aspirated Stimuli and Extralinguistic factors</p>	<ul style="list-style-type: none"> • SA program: No effect for SA program • Grammar test: No significant correlation between grammar score and Asp accuracy • NS contact: No effects of WeeklyContact and DailyContact on Asp accuracy • TL use: Significant positive correlations between Asp accuracy at Time 2 and (A) % use of Spanish overall, (B) % use of Spanish when writing, (C) speaking Spanish with NS of Spanish. Nearly significant correlations with (4) % use of Spanish when speaking, (E) speaking Spanish with NSs of Spanish, (F) time spent watching media in English (negative correlation) • TL use cont: Significant positive correlations between Time 2-Time 1 accuracy % change over time and (A) % use of Spanish when writing, (B) speaking Spanish with NS of Spanish

Table 69. Summary of the results of the lexical decision task (SA = Study Abroad, AH = At-home, TL = target language)

Task	Results
Lexical Decision Accuracy (Time 1)	<ul style="list-style-type: none"> • SEVILLE: No differences among all conditions (90%-98%). • NON-ASP: No differences among conditions (91%-98%) except Asp real words (77%). • SA1/SA2/AH: Asp real words lowest accuracy (54%/53%/22%), but all groups were accurate with CodaS/DIS real words. Not as good at rejecting non-words, but rejected Asp non-words more easily than other non-words.
Lexical Decision Accuracy (Time 2)	<ul style="list-style-type: none"> • SA1: Asp real words lowest accuracy (69%). Other conditions not really different from one another and fairly high accuracy. Most accurate for CodaS/DIS real words and Asp non-words • SA2: Asp real words lower accuracy (73%) than all but DIS and CodaS non-words. Most accurate for CodaS and DIS real words. • AH: Asp real word accuracy very low (15%). Highest accuracy Asp non-words, CodaS/DIS real words. More accurate for real words than non-words, except for Asp condition (real > non).
Lexical Decision Accuracy over time	<ul style="list-style-type: none"> • SA1: Significant positive change over time in most conditions, including Asp real, but not CodaS/DIS real words. • SA2: Significant positive change over time in most conditions, including Asp real, but not CodaS/DIS real & Asp non-words. • AH: Significant positive change over time for CodaS non-words only. Significant <u>decrease</u> over time for Asp real.
Lexical Decision and Extralinguistic factors	<ul style="list-style-type: none"> • SA program: No effect for SA program. • Grammar test: Significant positive correlation between grammar score and accuracy in every condition <u>except</u> for Asp real words • NS contact: No effects of WeeklyContact and DailyContact. • TL use: Significant positive correlations between Asp real word accuracy at Time 2 and (A) speaking English w/NS of Spanish, (B) time spent reading in Spanish. Significant negative correlation with media use in English. • TL use cont: significant positive correlation between Asp accuracy change over time (T1-T2) and speaking English with English speakers.
Lexical Decision Response Time	<ul style="list-style-type: none"> • SEVILLE: RTs for real words faster than for non-words in all conditions and Asp real RTs were not different from CodaS • NON-ASP: CodaS/DIS real words faster than non-words, but Asp real/non-words the same. Asp real slower than CodaS/DIS real • All 3 L2 groups: CodaS/DIS real words faster than non-words, but RTs for Asp real/non-words were not different. Asp real word RTs were slower than CodaS/DIS real words (higher processing cost for Asp real words than CodaS and DIS real words). Asp real words basically processed as non-words in terms of speed.
Lexical Decision RT over time	<ul style="list-style-type: none"> • SA1: faster in all conditions at Time 2, including Asp real words • SA2: faster at Time 2 for all conditions, including Asp real, but not for CodaS real words and Asp non-words. • AH: <u>not</u> faster at Time 2 for Asp and CodaS real words, but faster for all other conditions.

Chapter 5 – Discussion and Conclusions

The goal of this study was to investigate the perception of a dialect-specific phonological variant, word-internal /s/-aspiration preceding a voiceless stop in Andalusian Spanish, by L1 English-speaking learners of Spanish as a second language who were exposed to /s/-aspiration in a study abroad context and compare them to L2 learners who were not exposed to /s/-aspiration (i.e., at-home context). Additional factors that differentiated individual learners in the study abroad context were also explored in order to determine if they played a significant role in the learners' encoding and representation of /s/-aspiration over time. These factors were: 1) contact with native Spanish speakers, 2) use of Spanish by the L2 learners during the semester, 3) grammatical proficiency (i.e., grammar test score), and a comparison of learners in two different study abroad programs. Two native speaker groups who differed in their use of /s/-aspiration were also included to serve as a baseline for comparison to the L2 groups.

In order to meet these goals, two different speech perception tasks were administered. On the one hand, the participants' ability to categorize /s/-aspiration according to the correct orthographic representation of <s> was tested using a forced-choice identification task. On the other hand, their ability to access real words containing /s/-aspiration in the mental lexicon and to reject non-words, also comparing performance in the aspirated condition to a coda sibilant condition and distracter conditions, was tested using a lexical decision task. This task also included a response time component to analyze the speed with which they accessed words containing /s/-aspiration in the mental lexicon compared to words of other conditions, which represents the degree of processing cost associated with lexical access. The L2 participants also completed a language background questionnaire at Time 1, a language contact and use questionnaire at Time 2, a word familiarity survey, and a grammar test. The native speaker groups completed a background questionnaire and the grammar test. The results from both perception experiments, which were presented in Chapter 4, are now discussed in light of the research questions,

which are repeated below. The general discussion will be followed by a discussion of the implications for our understanding of SLA during study abroad, implications for study abroad program design, limitations of the study, suggestions for future research, and conclusions.

Research Questions

1. How is word-internal coda /s/ aspiration in Western Andalusian Spanish perceived by L2 learners?
 - a. Do L2 learners associate aspirated variants of coda /s/ with orthographic <s>? What are the patterns of association when aspirated variants are not identified as <s>?
 - b. Does the presence of an aspirated variant of coda /s/ hinder accuracy and/or response times for L2 learners' lexical access?
2. Does longitudinal exposure to Western Andalusian Spanish affect word-internal coda /s/ orthographic association and lexical access patterns (i.e., study abroad versus at-home)?
3. Do the following individual factors play a role in predicting orthographic association and lexical access patterns at the beginning and the end of one semester?
 - a. Grammatical proficiency
 - b. Contact with native speakers
 - c. Spanish language use by the learner
 - d. Study abroad program

The Perception of Word-Internal /s/-Aspiration in Western Andalusian Spanish by L2 Learners – Research Question #1

The first research question asks how L2 learners of Spanish who are L1 speakers of English perceive word internal /s/-aspiration in Western Andalusian Spanish, both in terms of their mapping of /s/-aspiration to orthographic representations (identification) and lexical access. Specifically, the first research question refers to the learners' initial perception of /s/-aspiration prior to exposure. The first part of the question asks whether L2 learners associate aspirated /s/ with orthographic <s> and what the patterns of association are when aspirated /s/ is not identified as <s> (Task 1). The second part of the question asks how /s/-aspiration affects learners' lexical access of words in terms of both the accuracy and speed of lexical access (Task 2). The findings of the forced-choice identification task will be discussed first, followed by the findings of the lexical decision task.

Initial Identification Patterns (Time 1)

Native Speakers. First, it is important to establish the native speaker baseline even though the NS data were not the primary focus of the current study. The SEVILLE group's identification results showed that they performed accurately in all conditions (74.5% to 100% accuracy). In fact, the only condition for which accuracy was lower than 96.1 percent was the aspirated condition (74.5%). When the aspirated items were not perceived as having a coda <s>, the primary response was <f>, which made up 20.8 percent of the responses. As discussed in Chapter Four, this result was primarily driven by eight native speakers from the SEVILLE group that responded incorrectly 50 percent or more of the time in the Asp condition, seven of whom chose the <f> response categorically or nearly categorically. Personal communication with some of the SEVILLE group participants after data collection suggested that one reason for this association was that the stimuli were non-words and some participants, though they perceived aspiration, did not know how to categorize it in non-words. This is not surprising since people are not used to hearing non-words, particularly non-words containing a phonological variant. Clearly,

this was not the majority of respondents, as this group correctly associated aspiration with <s> 74.5 percent of the time and the accuracy rate increased to 92.4 percent when those eight participants were not included. Another possibility that is related and was suggested by one or two participants is that social and contextual factors related to the use of /s/-aspiration may have influenced some NSs to avoid identifying /s/-aspiration as /s/ in a formal task because they may consider it ‘incorrect pronunciation’ according to formal rules whether they use /s/-aspiration themselves or not. Given that some of the SEVILLE participants had a background in Spanish language education, this seems plausible. Schmidt (2011) also found the same type of pattern for the Argentine (i.e., aspirating) group. Another hypothesis that she gave was that, according to some previous research, standard phonological variants can be easier to perceive and process than non-standard variants, even for native speakers of that dialect (e.g., Ranbom & Connine, 2007; Sumner & Samuel, 2009). Ranbom and Connine (2007) argued based on their findings that standard phonological variants often correspond to orthographic representations and thus appear to have stronger lexical representations than non-standard variants, which have representations that are more gradient. This means that processing the standard variant occurs in a more rapid and accurate manner. In the case of Andalusian /s/-aspiration, there are also multiple acoustic cues that work together to produce the aspiration, which is a more complex temporal and acoustic makeup than a sibilant. This could contribute to differences in the processing of [s] and /s/-aspiration among NSs of Andalusian Spanish. In all, though, the SEVILLE group’s results showed that the task worked as expected and that the majority of the time they associated aspiration with <s>.

The Non-Aspirating group, too, was very accurate, even more accurate than the SEVILLE group, in all conditions except the aspirated condition. This group’s rate of association of /s/-aspiration with <s> was very low (26.1%), demonstrating that native speakers that do not aspirate /s/ largely do not associate it with orthographic <s>, which is similar to the findings of Schmidt (2011) for a group of non-aspirating native speakers from Colombia. That speakers of certain dialects do not make use of acoustic

cues from other dialects of the same language has also been found in other perception studies (e.g., Scott & Cutler, 1984). There were, though, four individuals that associated /s/-aspiration with <s> between 29 percent and 70 percent of the time, which may have resulted from contact with aspirating NSs. All of the Non-Aspirating participants were employed at a large university and some of them had colleagues from parts of Andalusia. Therefore, it is expected that they had some degree of contact with them. Some of the participants indicated after data collection was complete that they understood that some people produce /s/ with aspiration, but that in their dialect they do not, so they did not mark it as <s>. However, this will not be explored further since the focus of the research questions and methodology of the current study is the L2 learners' perception of /s/-aspiration.

L2 Learners. The results from Time 1 for the L2 learner groups showed a low overall association of aspirated /s/ with <s> (0.3% to 18.1%), while at the same time showing significantly higher accuracy for the orthographic identification of all other coda conditions (79.2% to 100%). This indicates first that the L2 learners understood the task despite the presence of only non-words. They were able to associate the target CodaS, No Coda and control stimuli with the correct orthographic representations the majority of the time. In fact, the learners' identification of CodaS, No Coda and control stimuli was accurate over 90 percent of the time for all of these coda conditions except for [f], which still showed significantly higher accuracy than the aspirated condition. Therefore, we can interpret the significantly lower rate of <s> response in the aspirated condition as evidence that the learners in all three groups at Time 1 had not encoded /s/-aspiration as a variant of /s/ to an extent that would cause them to associate the acoustics of Andalusian /s/-aspiration with the grapheme <s>. This was also demonstrated by a wider range of variation in response types for the aspirated stimuli than for any other type of stimuli.

Response Types. When aspiration was not perceived as <s>, it was primarily perceived as if No Coda consonant were present, as hypothesized in Chapter 2. This was especially true for the At-home

group (95.8% No Coda responses), but it was also true of the Study Abroad 1 and Study Abroad 2 groups, which responded with the No Coda option 75.1 percent and 77.9 percent of the time respectively at Time 1. This preference for No Coda among the L2 groups at Time 1 demonstrates that breathy phonation or aspiration preceding the voiceless stop and/or the long-lag VOT were generally not parsed as acoustic cues for a variant of the phoneme /s/. The combination of the lack of encoding of these acoustic cues led to the perception of /CV.CV/ structure the majority of the time. This is in accord with the predictions of both the PAM and SLM outlined in Chapter 2, though they arrived at the same prediction differently. The predictions based on the PAM were that glottal activity preceding the voiceless stop would either be perceived as equivalent to English /h/, and thus not /s/ (if it was voiceless and strident), or perceived as part of the vowel if glottal activity was characterized by shorter breathy phonation. Additionally, the PAM predicted that long-lag stops would be mapped to English voiceless stop categories and therefore not parsed as an acoustic cue for /s/-aspiration. The L2 learners' data patterned according to this prediction. Though the SLM's predictions were different, they arrived at the same hypothesis. The SLM predicted that glottal activity preceding the stop would not be mapped to any English category because of differences in the constraints on word position in each language and that a new phonetic category would have to be created with time. Thus, the interpretation based on the SLM is that at Time 1 without previous exposure, no new category had yet been formed for /s/-aspiration because there had not yet been sufficient input, which resulted in the majority of responses reflecting the absence of a coda consonant. There was a difference, though, between the at-home and study abroad groups, in that the study abroad groups did seem to perceive *something* more often than the At-home group, which almost categorically chose No Coda. Thus, while At-home learners were "deaf" to the acoustic cues (Dupoux et al., 1997), a small number of the study abroad learners were more attuned to them, at least enough to perceive some consonant-like qualities.

Apart from No Coda responses and <s> responses to aspirated stimuli, <f> and <l> were the next most frequent responses given. However, <f> and <l> responses still only made up 2.5 percent to 6 percent of the responses combined for each group in this condition. Even fewer responses to aspirated stimuli were <r> and nasals. An explanation for mapping aspiration to <f> is that aspiration or breathy voicing preceding the voiceless stop and/or a long-lag VOT, which is a voiceless burst of air, have fricative-like noise qualities, at least to some degree, which may have led to the perception of the labiodental fricative. However, there may also be another explanation. The use of non-words is helpful in that it removes the problematic influence of word knowledge in an identification task, but it can also lead to other issues, as described previously based on the researcher's communication with a small number of the SEVILLE group participants who responded with <f> to most or all of the aspirated stimuli. Thus, it is possible that the L2 learners who responded with <f> to a small number of the aspirated stimuli (i.e., 25 <f> responses total among all three L2 groups) perceived the aspiration to some degree, but did not know how to categorize it. Finally, the explanation for the <l> is not quite clear. It could be that a small number of learners saw no suitable orthographic representation for what they perceived and selected at random. These 'other' responses, though, only made up very small percentage of the total responses for each L2 group at Time 1.

The coda condition apart from the aspirated condition that demonstrated the most variation in responses was the [f] condition, which resulted in 18 percent to 25 percent <s> responses for the three L2 groups. This result is most clearly explained by the acoustic similarities between [f] and [s], as both are voiceless fricatives that differ only in place of articulation. It is not surprising, then, that this coda condition exhibited the most variation apart from the aspirated condition. Schmidt (2011) also found some degree of variation in the [f] condition, but it was not to the extent that was observed in the current study.

L2 Group Differences. Next, it is necessary to discuss the group differences that were apparent even at Time 1 and possible explanations. While the two study abroad groups were not significantly different from one another in terms of the identification of /s/-aspiration, they were both significantly more accurate than the At-home group at Time 1. There are various potential explanations for this difference. First, the study abroad and at-home groups may not have been equal in terms of language proficiency and experience prior to participating in the study. In order to control for the At-home group having no prior study abroad experience, learners who had not yet taken any third year Spanish courses were recruited. There was less control over the prior experience of the study abroad learners in terms of the courses that they had taken prior to the study abroad period since they were all from different universities with different courses and course level systems. There were, then, no students in the At-home group that reported having taken three or more years of university-level Spanish courses. In the study abroad groups, though, 11 to 12 students, approximately half of each group, reported having taken three or more years of university Spanish. It seems, then, that there was a higher overall level of target language experience among the study abroad learners than the At-home group.

Another difference between the three L2 groups can be seen in the results of the grammar test at Time 1, which showed that the At-home group's Time 1 mean score ($M = 10.44$, $SD = 2.52$) was significantly lower than the Study Abroad 2 group's score ($M = 10.96$, $SD = 3.69$, $p < .05$), and both the Study Abroad 2 and At-home groups' scores were significantly lower than the Study Abroad 1 group's mean score ($M = 13.04$, $SD = 3.70$, $p < .001$), according to independent samples *t*-tests. Therefore, the At-home group's mean grammar score was significantly lower than that of both study abroad groups at Time 1. While the mean score was close to that of the Study Abroad 2 group, the At-home group showed a smaller standard deviation, indicating more students that were closer to the mean of 10.44 than the Study Abroad 2 group. This shows that there were more Study Abroad 2 group learners that achieved higher scores than At-home group learners. Though the factor of grammar test score will be

discussed in more detail later, it is important to note that there was a lack of a statistical effect for grammar test score at Time 1 and Time 2 in the aspirated condition, showing that grammar test score was not a good predictor of the identification of aspirated stimuli. Therefore, given the lack of predictive power for grammar test score in this study and that the results of the forced-choice identification task showed that the At-home group patterned like the two study abroad groups for all coda conditions except for the aspirated condition, we look to another potential explanation for the difference between the study abroad and At-home groups for the identification of aspirated stimuli.

First, it is possible that the learners who studied abroad had been exposed to more /s/-aspiration in general than the At-home group prior to the semester of data collection, and that this could have affected the initial categorization of /s/-aspiration by a small number of study abroad learners. On the Time 2 questionnaire, the learners were asked to report whether they had contact with any NSs from southern Spain or the Caribbean before arriving in Seville. This is because /s/-aspiration in Caribbean Spanish has some similar acoustic properties to /s/-aspiration in Andalusian Spanish. For the Study Abroad 1 group, two had reported knowing someone from southern Spain prior to arriving in Seville and one reported having watched YouTube videos of someone from Almeria before arriving in Seville. However, the two that reported knowing someone from southern Spain before arriving in Seville did not list anyone on the Time 1 questionnaire when asked to describe NSs of Spanish with whom they had contact for two years prior to studying abroad. Therefore, it is not clear what those relationships entailed. Seven participants in the Study Abroad 1 group reported having a friend or family member from Spanish speaking countries in the Caribbean (unspecified locations) prior to studying abroad and one reported that his parents had hired migrant workers from the Caribbean. Again, these reports were largely inconsistent with their responses on the Time 1 questionnaire. For the Study Abroad 2 group, seven participants reported having had a teacher from southern Spain prior to studying abroad in Seville and one reported having had an acquaintance from southern Spain. Also, 14 reported having a family

member (N=1), teacher (N=7) or friend (N=6) from a Spanish-speaking country in the Caribbean (unspecified locations). This group, too, generally did not report contact with these individuals on the Time 1 questionnaire and there is therefore no information on the relationships that the learners had with those individuals, or even whether they understood the geographic specifications of the question. For example, one individual thought that Fiji was in the Caribbean. It is possible that at least some of the seven participants who reported having had a teacher from southern Spain prior to studying abroad were not completely familiar with what constitutes southern Spain (i.e., Andalusia). It can be concluded, then, that even though the reports of the study abroad learners regarding their contact with NSs from the Caribbean and southern Spain conflicted between the Time 1 and Time 2 questionnaires, they did report, at least at Time 2, having been exposed to NSs of Spanish from the Caribbean and southern Spain prior to studying in Seville.

The question remains as to whether this influenced their Time 1 results on the forced-choice identification task. A comparison of the rates of correct <s> responses in the Asp condition to their reported contact with NSs showed that there was a range of scores in each study abroad group that did not show a pattern of advantage for those who reported contact with native speakers prior to studying abroad. There were learners who reported no contact with NSs and also had higher scores relative to others for /s/-aspiration, and there were also learners who reported contact with various native speakers who had lower scores for /s/-aspiration relative to the rest of the group. Therefore, there can be no argument made that prior contact with NSs, at least as reported on the questionnaire, gave an advantage to some learners at Time 1 on the identification task when responding to /s/-aspiration. It is possible, though, that prior knowledge of Spanish from other sources such as the classroom may also have allowed for the use of metalinguistic knowledge on the forced-choice identification task, which was an untimed task with low cognitive load.

It must be noted that, mistakenly, the At-home questionnaire did not ask the exact same question as the study abroad questionnaires. The At-home group was asked about their contact with NSs from southern Spain and the Caribbean *during* the semester of data collection, but not prior to it. Thus, their responses cannot be directly compared to those of the study abroad groups. Based on the Time 1 At-Home Questionnaire, though, there were three participants that reported contact with individuals from regions where /s/-aspiration occurs. There was one participant that reported annual contact with someone from Seville, one that reported daily contact with a teacher from Santiago, Chile, and one that reported monthly contact with a friend from Venezuela. Thus, these three people may have been exposed to /s/-aspiration prior to data collection. However, the strikingly low rate of <s> responses in the Asp condition for the At-home group at Time 1 makes it impossible to compare responses to their reported contact with NSs.

Another potential contributor to the difference between the study abroad and At-home groups is that the study abroad groups had been in Seville for two or three days (Study Abroad 1 group) to a week or week and a half (Study Abroad 2) prior to data collection. Consequently, they were most likely exposed to *sevillano* Spanish speakers to some degree prior to the Time 1 data collection. However, whatever exposure they had was not sufficient in such a short time span to lead to a high percentage, or even chance-level percentage of associations of /s/-aspiration with <s>. Yet, a small amount of exposure, especially to some of the more high frequency words, may have triggered the beginning of the acquisition process for a small number of learners that led to 16.7 percent (Study Abroad 2) and 18.1 percent (Study Abroad 1) correct responses to aspirated stimuli. Schmidt (2009), for instance, found that after just a three-week stay in the Dominican Republic, L2 learners were already increasing their comprehension of words and phrases that contained dialect-specific phonological variants. While this is clearly not indicative of a total phonological restructuring, it shows that it may be possible to detect

phonological variants in a very short period of time and begin the process of categorizing them, especially in words of high frequency that contain /s/-aspiration.

Finally, it must be acknowledged that the at-home and study abroad groups were different due to the fact that the learners who study abroad self-selected to study abroad. Consequently, there could be a degree of sampling bias according to learner characteristics. In other words, the learners in the study abroad groups chose to travel to another country in order to study Spanish. We do not yet know whether the learners in the At-home group will do the same, since they had not yet studied abroad. The learners who studied abroad may be a different type of language learners compared to the At-home group, have different goals, and participate in different types of language learning activities to reach those goals (Allen, 2010).

Summary. The results of the forced-choice identification task at Time 1 showed that L2 learners without extensive prior exposure to Andalusian /s/-aspiration, upon hearing isolated non-word stimuli containing /s/-aspiration prior to a voiceless stop of this variety, largely perceived /CV.CV/ rather than /CVs.CV/, as hypothesized. On the whole, they had not encoded /s/-aspiration as a variant of the phoneme /s/ to the extent that they would identify non-words containing it as containing a coda /s/. The SEVILLE group, though, did identify /s/-aspiration as a variant of /s/ for the most part. The Non-Aspirating group, on the other hand, was not much more accurate than the study abroad groups, showing that they did not categorize aspiration as <s> and also preferred the No Coda response.

Initial Lexical Decision Patterns (Time 1). The second part of the first research question asks how Andalusian /s/-aspiration initially affects L2 learners' lexical decision, both in terms of accuracy and response time. The hypothesis is that /s/-aspiration should lead to more inaccurate lexical decision responses compared to words containing a full coda sibilant, and that words with /s/-aspiration that are correctly accessed in the mental lexicon should demonstrate a greater processing cost, resulting in longer response times for words containing /s/-aspiration than words containing a sibilant [s].

Native Speakers. First, it is again important to establish the baseline of the SEVILLE group. Crucially, this group showed no significant difference in accuracy of lexical decision between CodaS and Asp real words. The SEVILLE participants were also equally fast at accessing Asp real words in the mental lexicon as they were CodaS real words. This is to be expected since they are native speakers of the target dialect. A significantly lower accuracy rate and response time for Asp real words compared to all other conditions among the SEVILLE participants would indicate a problem with the stimuli. However, the task clearly worked as it was designed to work for the SEVILLE baseline group.

Another notable finding was that the Non-Aspirating group performed like the SEVILLE group in all conditions except for the Asp real word condition, at least in terms of accuracy. In other words, the Non-Aspirating group's lexical decision was hindered to some degree by the presence of /s/-aspiration. The mean (76.61%, $SD = 20.60$) showed that the accuracy of lexical decision was still fairly high, as would be expected since they are native speakers with native speaker vocabularies and the words are real words, containing other phonological information apart from /s/-aspiration that will help trigger lexical access during word processing. However, a standard deviation that was twice that of the SEVILLE group and much larger than for any other condition indicates that this group's performance varied significantly more when /s/-aspiration was present than when it was not. Since some of the Non-Aspirating participants likely had contact with some Andalusian Spanish speakers is possible that the degree to which these participants had contact with Andalusian speakers or other /s/-aspirating speakers influenced the degree of variability in this group. This was not investigated in detail in the current study. Another factor could be the Non-Aspirating participants' place of origin and its proximity to Andalusia. Three of the Non-Aspirating participants were from Spain and thus much closer to Andalusia than those from Central and South America. The public awareness of how Andalusians speak and experience speaking with Andalusians must be higher in Spain than in these other countries due to media and other

forms of exposure, and this could be a source of variability among non-aspirating speakers. Again, this was not the focus of the current study and will not be explored beyond conjecture.

L2 Learners. Now that the baseline has been established and the task was shown to work correctly among the SEVILLE participants, the L2 findings will be discussed. Regarding the question of whether /s/-aspiration hindered the L2 participants' lexical access initially, the answer is clearly affirmative. This was evident both in terms of accuracy and response time. All three groups showed the lowest accuracy of all the conditions in the Asp real word condition. This was in contrast to the high level of accuracy found in the CodaS real word and Distracter real word conditions for all three groups, showing that among the real words, only aspirated real words caused lexical access problems for all three learner groups. Importantly, at the same time, the L2 learner groups were able to reject non-words containing aspirated /s/ at nearly the same rate that they successfully accepted real words containing coda [s] and the real word distracters. The combination of ease in rejecting aspirated non-words and difficulty in accepting the real aspirated words indicates an overall bias toward rejecting words that contained aspirated /s/, which is the opposite of the bias found for the CodaS and Distracter words. In the CodaS and Distracter conditions, all three L2 learner groups found it easy to accept the real words and more difficult to reject the non-words, showing a stronger tendency to accept a word as a real word if it did not contain aspirated /s/ than when it did.

It must be acknowledged, though, that another potential explanation for the observed difficulty in rejecting CodaS and Distracter non-words is the task design. Some non-words in the lexical decision task were created by taking one of the real words and changing a consonant, while for others a vowel was changed. These changes are outlined in detail in Table 29 of Chapter 3. This can lead to more difficulty rejecting some non-words than others because some vowels may sound more similar to each other than some consonant pairs and cause some non-words to sound more similar to their real word counterparts than others. An attempt was made to control for this effect as much as possible by first

trying to choose vowels for the non-words that were sufficiently different from the original vowel and second by eliminating any items from the analysis that caused significant problems for the SEVILLE group. It was clear that the reason the SEVILLE group had difficulty with those items was because of how the non-word was derived (e.g., baño [baño] vs. bañu [baɲu]). Thus, while the analysis was free of items that caused significant problems for the SEVILLE group, there may have been a residual effect in the L2 data, particularly for CodaS non-words and Distracter non-words, which exhibited significantly lower accuracy than Asp non-words generally. This will be discussed further in the discussion of the limitations of this study. Even if there was such an effect, though, the results still showed a clear pattern of differentiation for Asp words compared to the other conditions among L2 learners.

The response time data support the accuracy findings. While the SEVILLE group showed a consistent pattern across conditions of accepting real words faster than rejecting non-words, the L2 groups only followed this pattern for the CodaS and Distracter conditions. In the Asp condition at Time 1, the L2 groups' response times were similar between the real and non-words, and even slightly faster for rejecting non-words, though the difference was not significant. What this shows is that, while CodaS and Distracter real words were accessed faster than their non-word counterparts were rejected, as would be expected given the existence of lexical and phonological representations for real words and not for non-words, real and non-words that contained /s/-aspiration were accepted and rejected equally slowly and were processed at the slower speed of non-words rather than the faster speed of real words. Taken together, the accuracy and response time data for the L2 learners at Time 1 clearly show a higher processing cost and greater inaccuracy of lexical access for aspirated /s/ preceding a voiceless stop than for sibilant [s] preceding a voiceless stop and words that contained neither (i.e., distracters).

L2 Group Differences (Lexical Decision). While the three L2 groups patterned similarly in most conditions, it is again imperative to discuss the clear difference between the Study Abroad groups and the At-home group in the Asp real word condition in particular at Time 1. While the Study Abroad 1 and

2 groups' initial accuracy was in the low 50 percent range, the At-home group's accuracy was approximately 30 percent lower, in the low 20 percent range. As in the discussion of the results of the forced-choice identification task, there are various possible explanations for this pattern. First, there was a general trend of the At-home group to exhibit lower mean accuracy rates in nearly all of the conditions, which seems to reflect either a difference in experience with the target language between the Study Abroad and At-home groups or differences in learner characteristics among the groups, as was discussed regarding the results of the forced-choice identification task. However, the proficiency explanation is not sufficient in and of itself given that the difference in accuracy between the At-home and Study Abroad groups was relatively small for all of the conditions except for the Aspirated real word condition, where the difference was much larger. The results of the analysis of the grammar test scores at Time 1 compared to the learners' responses in the Aspirated real word condition also showed that grammatical proficiency was not correlated with their performance in that condition, whereas it was for all other conditions. The conclusion of that finding is that we cannot explain the group differences in the target condition based on grammatical proficiency. This was confirmed by a subsequent exploratory analysis that compared the accuracy rates in the Aspirated real word condition of the lexical decision task of the 10 learners in the At-home group with the highest grammar test scores at Time 1 to the accuracy rates of the 10 learners in the Study Abroad groups with the lowest grammar test scores at Time 1. This analysis, which was conducted comparing the grammar test scores and lexical decision scores in the Aspirated real word condition using the Mann-Whitney U test²², showed that despite a significantly higher median grammar test score for the top 10 learners in the At-home group ($Mdn = 12$, $SD = 1.85$) compared to the bottom 10 learners in the Study Abroad groups ($Mdn = 7$, $SD = 1.88$, $U = 288.00$, $p < .001$), the bottom 10 learners in the Study Abroad groups still had a significantly higher

²² A non-parametric t -test (Mann-Whitney U) was chosen because the group sizes were not equal and because the groups were a subset of the larger groups and thus smaller, the data were not normally distributed. The Mann-Whitney test is an ordinal test and it is typically preferable to report the median.

accuracy rate in the Aspirated condition ($Mdn = 57.5\%$, $M = 57.9\%$, $SD = 17.6$) than the top 10 learners in the At-home group ($Mdn = 33.2\%$, $M = 16.3\%$, $SD = 12.0$) at Time 1 ($U = 2.5$, $p < .001$). Thus, we can conclude that grammatical proficiency as measured by the grammar test is not a good predictor of being able to access words from the mental lexicon containing /s/-aspiration at Time 1.

Since grammatical proficiency is not a good predictor, it is necessary to search for other possible explanations for the observed group differences. One other possible explanation for such a large difference is prior exposure to native speakers of Spanish, particularly those from southern Spain or the Caribbean, where /s/-aspiration is common and shares similar acoustic properties. In order to investigate the question of prior exposure, the data from the questionnaire that contained learners' self-reported contact with native speakers from southern Spain and the Caribbean prior to data collection was compared to their scores at Time 1 in the Aspirated real word condition of the lexical decision task. As was seen in the discussion of the results of the identification task, there was again no clear connection between having prior exposure to native speakers of Spanish from the Caribbean or southern Spain and having a higher initial rate of accuracy in the Aspirated real word condition of the lexical decision task. There was a wide range of scores in this condition at Time 1 and some learners who reported contact had very low scores and others who reported contact had higher scores. There were also learners who reported no contact and still had higher scores. Therefore, in the current dataset there is no justification for arguing for an effect of prior exposure. That is not to say that it played no role in their performance. But further data with more detail on their prior exposure would be needed to be able to draw more solid conclusions.

Another possible explanation is that the learners who studied abroad had received some level of input prior to data collection between the time of their arrival and the day of their participation. This ranged from approximately two days (Study Abroad 1 group) to one or one and a half weeks (Study Abroad 2 group). The learners in the Study Abroad 1 program participated two to three days after arrival

and the learners in the Study Abroad 2 program participated one to one and a half weeks after arrival. It is possible that their exposure to Andalusian Spanish, though only for a short time, influenced their performance on the lexical decision task and helped them to detect /s/-aspiration in the input and respond at a near-chance level prior to converting the input to intake and beginning the process of encoding aspirated /s/ as a variant of /s/ in the L2 phonological system. This could be particularly true of high frequency lexical items that contain /s/-aspiration, but there is not sufficient data to investigate this hypothesis.

Given that the previous explanations are insufficient to explain the observed differences between the Study Abroad and At-home groups, a more likely hypothesis is that there was some degree of sampling difference according to learner characteristics. In other words, the learners who studied abroad may generally be of a different type of language learner, with different learning goals, than the learners in the At-home group. This could, in turn, affect outcomes (Allen, 2010). They did, after all, self-select to study abroad. They did not have to go abroad to study the foreign language, but chose to do so. It could be that many of the learners in the At-home group will never study abroad and are primarily learning Spanish to fulfill a requirement.

In order to further investigate sampling differences between the groups and demonstrate whether the patterns that obtained for unequal groups (i.e., the entire Study Abroad and At-home groups), an analysis was conducted to compare equivalent subsets of the Study Abroad and At-home groups in terms of their performance in the Aspirated real word condition at Time 1. The 10 learners in the At-home group with the highest scores in the Aspirated real word condition of the lexical decision task at Time 1 were compared to the 20 learners in the Study Abroad groups (i.e., 10 from each group combined into one 20-person Study Abroad group) with the lowest scores in the Aspirated real word condition at Time 1. The At-home group had a mean score of 36.18 percent (*Mdn* = 31.58, *SD* = 18.27) and the group of 20 study abroad students had a mean score of 32.91 percent (*Mdn* = 35.00, *SD* =

13.32). A Mann-Whitney U test showed that there was no significant difference between the two groups ($U = 97.50, p = .912$). This means that approximately the top half of the At-Home group was performing like the bottom half of the Study Abroad groups at Time 1 when responding to /s/-aspiration on the lexical decision task. This finding supports the sampling hypothesis because it shows that subsets of the two groups are comparable, but that in order to compare them it is necessary to take the highest performers from one group and the lowest performers from the other. Despite this difference in the groups, it is still possible to draw important conclusions and find clear differences among comparable subsets, as will be discussed in the following section about change over time on the lexical decision task.

Summary. The results of the lexical decision task at Time 1 showed clearly that the task worked as it was designed to work because the SEVILLE group's accuracy did not differ depending on condition. Also, this group showed continuity between conditions for response time, in that correct responses to real words were always accessed faster than non-words were rejected and response times for aspirated words were no different from words with a coda sibilant. The Non-Aspirating group, too, was very accurate overall and significantly less so in the Asp real word condition. Also, response times showed a greater processing cost in this condition for the Non-Aspirating group compared to other conditions. The L2 groups showed the greatest inaccuracy in the Asp real word condition, being significantly less accurate than both NS groups, and their accuracy for Asp real words was significantly lower than for all other conditions. The At-home group was also much less accurate than both Study Abroad groups, which patterned very similarly. However, it was shown that the highest performing at-home learners and the lowest-performing study abroad learners were equivalent at Time 1. Regarding response time, the L2 groups' patterns for were similar and showed a significantly greater processing cost for aspirated stimuli than sibilant and distracter stimuli. It can be concluded, then, that word-internal /s/-aspiration in Andalusian Spanish preceding a voiceless stop makes it significantly more difficult, both in terms of

accuracy and processing time, for L2 learners with no prior exposure to access words in the mental lexicon than when the same words contain a full sibilant.

The Effect of Longitudinal Exposure to Western Andalusian Spanish on Orthographic Association and Lexical Access Patterns – Research Question #2

Now that it is established that words in Western Andalusian Spanish that contain /s/-aspiration are more difficult for the L2 learners to phonologically encode and access from the mental lexicon than the standard variant ([s]), the second research question asks whether longitudinal exposure to Western Andalusian Spanish during a semester of study abroad positively affects how L2 learners perceive /s/-aspiration in terms of orthographic identification and lexical access. To answer this research question, the current study compared a group of L2 learners who stayed in the United States to learners that were enrolled in two different study abroad programs in Seville, Spain during the same time period of 10 to 11 weeks. The data were collected at the beginning of the semester and near the end of the semester, and the statistical analysis compared the results according to data collection time to determine whether there was a change over time within each group and a difference between the At-home and Study Abroad groups for identification and lexical access.

Orthographic Association over Time. The results of the forced-choice identification task showed that the learners in the Study Abroad 1 and Study Abroad 2 groups significantly increased their association of /s/-aspiration with <s> from Time 1 to Time 2 (10 to 11 weeks later) by 17.7 percent and 21.5 percent respectively, while the At-home group did not and still exhibited a very low rate of identifying aspiration as <s> at Time 2 (0.9%). It is important to consider, too, that the study abroad groups' rates of identification of aspiration as <s> even at Time 2 were nowhere near categorical, at 35.8 percent and 38.2 percent, and therefore did not reach the SEVILLE group's rate of 74.5 percent. Nevertheless, they were unmistakably moving in the direction of the NS norm for Seville, doubling their

accuracy rates for the identification of aspiration as <s> and reaching nearly half of the SEVILLE group's mean accuracy by Time 2. The largest change in orthographic mapping over time was a decrease in the selection of No Coda and an increase in the selection of <s> among study abroad learners, but not among at-home learners. In all the other conditions over time, the three groups performed similarly and were very accurate at both data collection times. The [f] condition exhibited the least accuracy at both times of all control conditions at both time points. The one group difference over time apart from the aspirated condition is that the Study Abroad 2 and At-home groups did significantly increase accuracy over time in the [f] condition while the Study Abroad 1 group did not. The Study Abroad 1 group did increase accuracy in this condition at Time 2 but it was not a significant increase.

There are a few different possible explanations for why the study abroad groups did not achieve a higher accuracy rate when responding to aspirated stimuli at Time 2 than they did. First, it is likely that three months was not long enough to see a complete phonological restructuring and native-like performance on perception tasks. This is a very plausible explanation because it has been found in some of the previous research on the acquisition of variable forms in a second language that L2 learners move in the direction of NS norms without quite reaching them, at least during the time window of the research study. Given the overall upward trajectory of acquisition that is evident in this and some previous studies on the acquisition of L2 variation (see Geeslin, 2011b for a review), it is reasonable to conclude that phonological encoding was beginning to take place and to hypothesize that more time in the target language community before re-testing would result in higher accuracy rates for the identification of /s/-aspiration. There was quite a lot of variation among the study abroad learners even at Time 2, as demonstrated by standard deviations and 95 percent confidence intervals that were much larger than for the other conditions. Additionally, the amount of variation among study abroad learners at Time 2 was greater than at Time 1 despite the higher overall mean. Previous SLA research has shown that as the interlanguage grammar incorporates new features (i.e., restructuring), variability increases

and as restructuring proceeds and nears completion, variability decreases. Thus, the study abroad learners in the current study seemed to be in the early stages of acquisition at Time 2 for the identification of /s/-aspiration as a variant of /s/ as represented by orthography.

A second explanation, which likely exists in tandem with the first, is that the identification task was based completely on non-words. As argued before, this could have negatively affected the L2 learners' performance on the task. This is likely since the study abroad groups' accuracy rates for aspirated items on the forced-choice identification task at Time 2 were quite a bit lower than accuracy for real aspirated words in the lexical decision task, while the control conditions of the forced-choice identification task exhibited very high accuracy overall. This is precisely why two tasks, which measured two different aspects of the perception of /s/-aspiration, were incorporated into the methodology. The lexical decision task, which did include half real words with the non-words and tested a different aspect of perception (i.e., lexical access), showed higher accuracy in the aspirated real word condition for all groups than the forced-choice identification task. The results of the lexical decision task, then, mediate the somewhat unexpected results in the aspirated condition of the forced-choice identification task, particularly for the SEVILLE group, by showing what occurs when the participants are presented with real lexical items that have lexical and phonological representations as opposed to non-words that do not.

The results of the forced-choice identification task, though, do clearly show that studying abroad in Seville for one semester had a significant positive effect on L2 learners' identification of /s/-aspiration over time, while no such effect was observed for the L2 learners who were not exposed to Western Andalusian /s/-aspiration in the traditional foreign language classroom environment at their home institution. This is an expected result, since a lack of input logically leads to a lack of acquisition. In fact, the opposite finding would have been problematic because it would have meant that the task allowed learners who did not have /s/-aspiration in the input to perform more accurately the second time

around, and any observed increase in accuracy could have then been considered a mere task effect. Clearly, this was not the case in the current study.

As for what the results tell us about the study abroad groups' L2 phonological systems, while orthographic and phonological representations are not always one-to-one and we must be careful in interpreting orthographic mappings too directly with phonological representations in every case, the results are indicative of a pattern of beginning to categorize /s/-aspiration as a legitimate allophonic variant of the phoneme /s/ in Andalusian Spanish. While English and Spanish have differences in their allophones of the phoneme /s/, the orthography-phonology relationships for Spanish and English /s/ are the same for the sibilant allophone of this phoneme. In other words, both languages map the voiceless alveolar fricative to orthographic <s>. This is demonstrated in the L2 learners' performance in the CodaS condition, which was 99 percent to 100 percent correct for all three L2 groups. Knowing this we can assume that if L2 learners choose a response that includes orthographic <s> in coda position, it means that they perceived something that represented the phoneme /s/ as opposed to /f/ or something else, such as No Coda. Applying this to the aspirated condition, the results over time for the study abroad groups indicate phonological restructuring that was beginning to encode /s/-aspiration as a legitimate allophone of the phoneme /s/.

Lexical Access over Time. Like the results of the forced-choice identification task, the lexical decision task showed that the two study abroad groups significantly increased their accuracy of lexical access and speeded their processing of Asp real words from Time 1, at the beginning of the semester, to Time 2, near the end of the semester abroad. These two groups were not significantly different in terms of change over time in the accuracy and response time in the Asp real word condition. In fact, both study abroad groups' mean lexical decision accuracy rates were high enough at Time 2 that they were no longer significantly lower than that of the Non-Aspirating group, reaching approximately 70 percent accuracy for aspirated words. This indicates that the learners, through exposure to /s/-aspiration, were

able to access words containing /s/-aspiration in the mental lexicon to a degree that was similar to native speakers of Spanish that do not aspirate /s/. The study abroad groups' response times for aspirated stimuli were also like the Non-Aspirating group's RTs. While this is not equivalent to having performed like the native speakers of the SEVILLE group, it is a noteworthy achievement since lexical decision tasks are difficult due to the speed at which participants must respond. The results confirm, then, that being exposed to a phonological variant leads to more accurate and faster lexical access of words that are produced with the target phonological variant, even in just under three months' time.

At the same time, the At-home group actually decreased lexical decision accuracy over time and did not show faster processing of aspirated real words at Time 2 than Time 1, even though at the same time they did show faster processing in four of the six conditions of the lexical decision task at Time 2. This result confirms that not being exposed to /s/-aspiration in the input does not help learners to better access aspirated words in the lexicon, and it follows the foundational principle of SLA regarding the necessity of input for acquisition to occur. In terms of the overall lexical decision results for the At-home group in all conditions, there was a pattern over time of decreasing accuracy in all three real word conditions and increasing accuracy in all three non-word conditions, though only one of these differences over time was statistically significant (CodaS non-words). This pattern is indicative of a slight tendency at Time 2 toward rejecting stimuli as non-words regardless of condition. Importantly, the results of the At-home group in the aspirated condition cannot be completely attributed to a lower level of proficiency. Though there were some differences in terms of grammar test scores and history of Spanish instruction between the at-home and study abroad groups, the At-home group showed accuracy rates in the CodaS real word/non-word and Distracter real word/non-word conditions that, though significantly less accurate than the study abroad groups, were much closer the two study abroad groups' mean accuracy rates than the At-home group's accuracy rate in the aspirated condition was to the study abroad groups' accuracy in the aspirated condition. Figure 15, comparing Time 1 and Time 2

for the three L2 groups and repeated below as Figure 22 for the reader's convenience, shows unmistakably that, of all the comparisons between the at-home and study abroad groups, the aspirated real word condition was by far the largest difference between the at-home and study abroad groups, even at Time 2. This shows that /s/-aspiration was an exception to the general patterns and was the specific condition that posed the greatest difficulty for at-home learners and the greatest difference between this group and those who studied abroad.

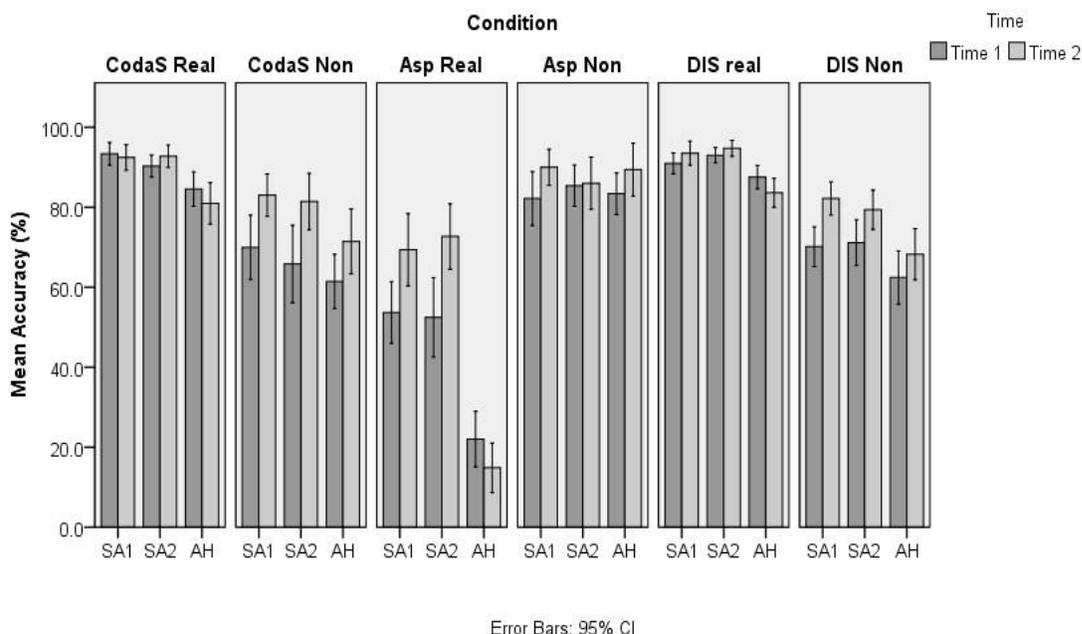


Figure 22. Summary of lexical decision accuracy results for the L2 groups according to Time and Condition (N=73)

It is crucial, though, to account for the finding that there is a sampling difference between the groups and to ensure that the results of the current study that show positive change over time for the Study Abroad groups but not for the At-home group still hold when comparing equivalent subsets of the data. Taking the two subgroups, the top 10 performers of the At-home group at Time 1 and the bottom 10 performers in each Study Abroad group at Time 1 (i.e., 20 study abroad learners total conflated in one group), a Mann-Whitney U test was run on the Time 2 results for the Aspirated real word condition of the lexical decision task to compare the two groups. The test showed that at Time 2 the Study Abroad

sub-group had a significantly higher accuracy rate ($Mdn = 74.3\%$, $M = 67.2\%$, $SD = 23.8$) than the At-home sub-group ($Mdn = 12.5\%$, $M = 12.8\%$, $SD = 11.8$, $U = 4.5$, $p < .001$). Compared to Time 1, the At-home sub-group decreased their rate of correct responses over time, while the Study Abroad sub-group increased correct responses by 34.3 percent. This shows that, when comparing subsets of the At-home and Study Abroad groups that were statistically equivalent at Time 1 in the Aspirated real word condition of the lexical decision task, positive change over time was still only observed for those who were in the study abroad learning context and received input that contained /s/-aspiration.

Summary. Tarone (2007) argued that the only real way to determine whether social context plays a role in the development of the interlanguage is to conduct studies that follow learners' development of specific L2 forms over a period of time as they interact with interlocutors in specific social contexts, and that making use of the study abroad context in L2 research is a promising way of doing so (p. 845). The results of the current study, which tracked study abroad learners' development of the encoding and lexical processing of a geographic and sociolinguistic variant over time, has shown that learners' immersion in the geographic context of Seville, Spain, where /s/-aspiration is prevalent in the input, changed their interlanguage phonological system regarding the encoding of /s/-aspiration as a legitimate variant of the phoneme /s/ and the ability to access words containing the variant from the mental lexicon, and that not being exposed to the variant did not lead to such change. Critically, this is just one aspect of the acquisition of sociolinguistic competence during study abroad. It is not argued here that these learners were acquiring the social characteristics of /s/-aspiration through interactions with native speakers in different social contexts. Future research is necessary in order to target different social contexts and learners' interactions with different interlocutors in order to understand whether learners acquire the social values related to /s/-aspiration in Seville during SA. However, the current study has taken a necessary first step in showing that the geographic context in which L2 learners acquire Spanish as a second language has a positive impact over time on the development of their L2

phonological systems with respect to their encoding and processing of a dialect-specific phonological variant.

Given that there was significant positive change over time in the aspirated condition for the study abroad learners, which reflects their exposure to /s/-aspiration in the input in a specific geographic context, it is necessary to break down the results according to the extralinguistic factors that were analyzed in order to determine what effect they had on the results of the individual learners within the study abroad groups.

The Effects of Extralinguistic Factors on Orthographic Association and Lexical Access Patterns – Research Question #3

The final research question seeks to determine the effects of the following extralinguistic factors on study abroad learners' perception of /s/-aspiration over time: study abroad program, grammar test score, language contact with NSs, and target language use. These factors were operationalized by means of a 20-item grammar test and a language contact and use questionnaire that was administered at Time 2. The results for the factor of study abroad program were observed through the statistical analyses of each task comparing the two study abroad groups over time. The first extralinguistic factor to be discussed is the learners' study abroad program.

Study Abroad Program. In order to determine whether learners enrolled in two different study abroad programs performed differently for the perception of /s/-aspiration over time, the results of the statistical analyses of each task comparing the L2 learner groups provide the answer. A significant difference between the learners in the two programs in performance on the perception tasks would indicate, perhaps, a difference in the structure of the study abroad programs that led to more opportunities for meaningful L2 input for one group than the other. Some previous research has found that different types of study abroad programs, such as those that focus on service learning versus those

that do not have such a focus, or programs located in certain locations that provide access to certain types of communicative situations, can lead to different amounts of target language use and opportunities for input (e.g., Dewey, Ring et al., 2013; Martinsen et al., 2010).

Ultimately, it was determined that the two study abroad groups began and ended the semester with no real differences between them. In other words, both showed high accuracy in all conditions except the aspirated condition initially and made very similar gains over time in the aspirated condition and other conditions. The reason for this is probably, at least in part, that the two study abroad programs, though they had differences in terms of program size and the number and content of course offerings, were of a similar structure. In other words, they were not inherently different types of programs and both were what would be considered 'traditional' study abroad programs. How 'traditional' is defined here is by a combination of factors. Both programs were one semester in duration. Both provided opportunities for the learners to engage with NSs, such as service learning, volunteer work, and language exchanges, but did not require them as part of the program design. Learners in both programs were housed with local host families in Seville and had a roommate that was a native English speaker. Both programs offered classes only in Spanish to the learners that participated in the study. The professors employed by both programs were native speakers of Spanish. The learners in both programs took classes alongside other native English speakers. Both programs strongly encouraged the use of Spanish, at least on the school premises, though there were some differences in how they did so. Study Abroad 1 made learners sign a language pledge and Study Abroad 2 hired local Spaniards to spend time at the school speaking in Spanish with students between classes and take them on frequent tours around the city to encourage the use of Spanish. However, learners in both programs were still observed speaking in English frequently during the researcher's visit. The learners within each program were largely of similar age, ethnic and educational backgrounds and the populations of learners in both programs and in the current study were predominantly female. The list of similarities

could go on, and more differences could likely be found upon closer inspection. But the main point is that, with so much overlap in how the two study abroad programs were structured and the types of opportunities that learners had to engage meaningfully with NSs of the target language, it is not surprising that there were no significant differences between the two groups. Perhaps if one of the groups were of the traditional type described above and one only took classes at the local university alongside NSs, there may have been more of a difference. It will be left to future research to investigate such comparisons in more detail.

Grammar Test Score. The next extralinguistic factor is the participants' grammar test scores at Time 1 and Time 2, as well as the change over time in grammar test score. The correlation analysis for grammar test score and performance on the perception measures produced a few interesting results. First, accuracy at Time 1 in the aspirated condition of each task was tested for correlation with initial grammar test scores in order to determine whether higher initial grammatical proficiency, as measured by the 20-item grammar test, was positively correlated with the identification of /s/-aspiration and lexical access patterns. Recall that the two study abroad groups' grammar test scores at Time 1 were significantly different and the Study Abroad 1 group's score was higher than the Study Abroad 2 group's score. The results of the correlation analysis for both tasks showed no significant correlation between Time 1 grammar score and Time 1 accuracy for the aspirated items of both tasks.

An interesting finding, though, was that the accuracy of the lexical decision task for every condition except the aspirated real word condition was positively correlated with Time 1 grammar score to a significant degree, showing that the study abroad learners who had a higher initial grammar test score also showed more accurate lexical access in all conditions except for the aspirated condition. This means that having a higher grammatical proficiency, at least as measured by the grammar test, was of a general benefit but did not help them when it came to the lexical processing of /s/-aspiration. Interestingly, this difference in effect of grammar score depending on condition was only observed for

all conditions in the task that included real lexical items (lexical decision task) and not for all conditions of the forced-choice identification task. A potential explanation for this finding is that learners with higher grammatical proficiency, and probably higher L2 proficiency in general, have more established L2 lexicons and more developed lexical processing abilities, which in turn facilitates faster and more accurate lexical access in general. This is in agreement with previous research that has shown that having a higher L2 proficiency has been linked with positive changes in neural structures, resulting in increased attention control (Segalowitz & Hulstijn, 2005) and cognitive control (e.g., Abutalebi & Green, 2007; Bialystok & Craik, 2010)²³. For example, Abutalebi and Green (2007) found a difference in how lexical retrieval occurred among low and high L2 proficiency bilinguals, showing that L2 learners with a higher L2 proficiency demonstrated more automatic processing and did not use the prefrontal cortex of the brain, while those with a lower level of proficiency did employ the prefrontal cortex, resulting in more controlled, non-automatic processing.

Interestingly, though, when participants heard words that contained the target variant (/s/-aspiration), the advantage of higher grammatical proficiency disappeared. In other words, the playing field was level because lexical access was more difficult for all, regardless of grammatical proficiency level, when an optional phonological variant was present in the aural stimulus than when it was not. This demonstrates that there are other forces at work apart from L2 proficiency and cognitive factors. Specifically, given the results of the current study, it is clear that learners need input that contains the target variant in order to acquire it, and when groups of learners are exposed to the variant in the input over a period of time, in this case approximately three months, there is a general positive effect for most of the learners that overrides differences based on L2 proficiency.

²³ See also Van Hell and Tanner (2012) for a review of research on the relationship between L2 proficiency and lexical activation

Cognitively speaking, there is an interesting avenue for future research to determine the cognitive processes that learners use to process input that contains an optional phonological variant. Although the methodology of the current study does not allow for making claims regarding cognitive processes, it seems plausible, based on the results of the current study, that L2 learners may have to employ different cognitive processes for non-standard variants compared to standard variants and may show different patterns in studies that use neural imaging technology. Based on the findings of previous research, it may be that L2 learners with higher proficiency do not employ the prefrontal cortex for the processing of standard variants and that their processing of these variants is automatic, but that non-standard variants require more 'active-controlled retrieval' (Petrides, 1998), leading them to employ the prefrontal cortex in a way that is similar to learners with lower levels of proficiency. Future research using technology such as fMRI would be able to shed light on whether such differences exist, and this is a methodological step that has been advocated by others for the study of sociolinguistic variation (e.g., Campbell-Kibler, 2010).

The next question related to grammar test score was whether the Time 2 grammar test score was significantly correlated with Time 2 accuracy. At Time 2, the Study Abroad 1 group's grammar score was still significantly higher than the Study Abroad 2 group's score even though each group's score increased. For the lexical decision task, the result was similar to Time 1. A higher grammar test score at Time 2 correlated positively with accuracy in the CodaS non-words and Distracter non-word conditions, and was approaching significance for the other CodaS real words, Distracter real words and Asp non-words. However, the study abroad learners' accuracy in the Asp real word condition did not correlate at all with Time 2 grammar test score. This, again, suggests that there was a general advantage of having a higher grammatical proficiency for the accuracy of lexical access. Interestingly, this may also be tied to research that has found that learning the L2 while immersed in the target language context leads to cognitive changes resulting in more automatic lexical processing and less reliance on the prefrontal

cortex, again showing an advantage in lexical processing for L2 learners of higher levels of proficiency (e.g., Stein, Federspiel, Koenig, Wirth, Lehman, Wiest, Strik, Brandeis & Dierks, 2009). Clearly, though, even at Time 2 the advantage was non-existent when presented with /s/-aspiration.

Finally, when the change over time in the grammar test score was tested for correlation with change over time of the accuracy in the aspirated conditions for both perception tasks, the results showed no significant correlations. This means that, in general, making grammatical gains over time was not related to the learners' ability to identify and access lexical items containing /s/-aspiration.

To summarize, the results for the effect of grammatical proficiency, as tested by means of the grammar test score, showed that one's grammatical proficiency either prior to study abroad or after study abroad did not help when it came to the perception of /s/-aspiration, despite an advantage in other conditions, at least for lexical access. What these results suggest is that, while lexical access seems to be generally more accurate for those of higher language proficiency, which likely reflects a more developed L2 lexicon, L2 phonological system, and greater L2 cognitive processing abilities in general, other factors are also at work apart from grammatical proficiency when it comes to the perception of optional phonological variants. It is clear based on the results of the current study that being able to encode and process an optional variant was not dependent on grammatical proficiency, at least in terms of the results of the grammar test. Future research may find that other cognitive characteristics are related to the encoding and lexical processing of optional phonological variants. However, the results of the current study suggest that it is necessary to discuss the effects of other factors that are related to the input that L2 learners receive in certain contexts, such as language contact and target language use.

Contact with Native Speakers. Contact with native speakers and target language use are strongly connected since interacting with NSs often entails target language use. First, language contact with NSs was operationalized as the number of NSs with whom the participants had 15-20 minute conversations on a weekly and daily basis. These numbers were then tested for correlation with the learners' accuracy

in each condition of each task. The results of the analysis showed that there were no significant correlations between the number of reported NS contacts on a weekly and daily basis and accuracy in any condition for either perception task. While on the surface this may seem to be a surprising result, it is not when exploring it deeper, and it is actually in accord with previous findings about language acquisition in the study abroad context.

Segalowitz et al. (2004) argued that it is not necessarily the number of contacts that L2 learners have with native speakers that leads to acquisition, but rather the nature of their interactions with NSs. With this in mind, there are a few important considerations regarding the data from the current study. First, the WeeklyContact and DailyContact factors did not include information regarding the characteristics of the interlocutors with whom the learners had contact and the types of input they received. It is very possible, then, that the reported numbers include NSs of Spanish from other dialect regions of the Spanish-speaking world, or even Andalusians who did not aspirate /s/ in conversations with the L2 learners. Communication Accommodation Theory (Giles, Coupland & Coupland, 1991) would suggest that native speakers may accommodate to L2 learners, simplifying their speech in any number of ways, in order to minimize miscommunication. This could result in more formal speech and more use of the standard variant [s] when speaking to less proficient students, or, alternatively, more use of /s/-aspiration with more proficient students. For this reason, we cannot assume that the lack of a statistical effect for WeeklyContact and DailyContact means that contact with NSs played no role at all in the learners' perception of /s/-aspiration over time. It simply means that these variables did not account for factors such as speech accommodation and the actual content of the input learners received in terms of the prevalence of /s/-aspiration. For this reason, language contact was also operationalized alongside target language use by means of the part of the questionnaire that asked learners to report their time spent engaging in the target language in various ways during an average 16-hour day.

Target Language Use. The use of the target language by the study abroad learners was determined by eliciting information about the following aspects of the study abroad learners' target language use: the percent use of Spanish overall during the semester, as well as specifically for writing, reading, and speaking; and the time spent speaking Spanish with NSs of Spanish and English, speaking English with NSs of Spanish and English, watching media in Spanish and English, and reading in Spanish and English during an average 16-hour day. These factors were tested for correlation with both Time 1 and Time 2 accuracy for each perception task, as well as the percent change in accuracy from Time 1 to Time 2. The findings for each task will be discussed separately.

Target Language Use and Orthographic Association. The results of the correlation analysis for the identification task showed there were three significant factors and two that were approaching significance. The strongest effect that was observed was a positive correlation between the time spent writing in Spanish and accuracy at Time 2 in the Asp condition. This means that learners who reported writing more in Spanish than in English during the semester also demonstrated higher identification accuracy for the aspirated condition at Time 2. At first, this result seems surprising because writing is a very different language skill than speech perception. However, there is a hypothesis that could potentially explain this finding.

It could be that L2 learners who spend more time writing in Spanish have different learner characteristics than those who spend less time writing in the target language. If they are more advanced and/or have different characteristics than those that do not write as much in Spanish, they may also seek out more opportunities to interact meaningfully with NSs and/or be more attuned to phonological variation than others. This would be in accord with research to date on the acquisition of variation, which has found that advanced learners tend to acquire variants more often than intermediate and beginner learners (Geeslin, 2011b). In other words, it could be that the percent use of Spanish while writing is a way of understanding different learner characteristics, such as those who set more specific

language learning goals during SA. Allen's (2010) findings support the potential explanatory power of this hypothesis for the observed patterns. She studied the reasons that learners studied abroad and their relationship to goal-setting and the activities in which they engaged during study abroad. She found that L2 learners studied French either for linguistic reasons or career-oriented reasons. Likewise, they saw studying abroad as either a step to achieving fluency in the L2 or a way of exploring the world through travel and cultural learning. Their reasons for studying French and studying abroad, in turn, affected the activities in which they engaged while abroad and the goals that they set for themselves. For some, this resulted in more effective goal-setting and taking responsibility for their own learning, while for others it resulted in shifting the blame for limited interaction with NSs of French on others. Those that set specific goals such as holding at least 20- or 30-minute conversations or being comprehensible in service encounters engaged in more specific language-related activities, while those who set generic goals such as improving speaking abilities or accent did not. The application of these findings to the current study is that writing more in the L2 may be an activity that results from specific goal-setting, which results from studying abroad for language-learning reasons as opposed to other reasons. This type of language learner, then, may be more likely to acquire variable forms than other learners that do not have similar characteristics.

The other two language use factors that correlated significantly with Time 2 identification accuracy in the aspirated condition of the forced-choice identification task were the amount of time that the learners reported speaking Spanish with NSs of Spanish on an average day and the reported percent use of Spanish overall during the semester. Both correlations were positive, and the two measures of Spanish language use were correlated with one another, showing that those who reported using Spanish more overall (i.e., in any language domain – e.g., writing, speaking, reading combined) and spending more time speaking Spanish with NSs of Spanish also demonstrated higher accuracy for the identification of /s/-aspiration as a variant of /s/. Interestingly, there was also a significant positive

correlation between the percent use of Spanish while writing and both the time spent using Spanish with NSs of Spanish and the percent use of Spanish overall. This shows that those who reported spending more time speaking with NSs and using more Spanish overall also spent more time writing in Spanish. These three factors taken together indicate that there were some learners who put more time and energy into speaking and writing in Spanish than others, and that those who did also identified /s/-aspiration as /s/ more often than those who spent less time writing and speaking in Spanish. Also, when the participants' percent change over time of accuracy in the aspirated condition was tested for correlation with the same factors, it was significantly and positively correlated with the percent use of Spanish when writing and the time spent speaking with NSs of Spanish in Spanish. Given that Time 2 accuracy and the Time 2-Time 1 percent change in accuracy were highly correlated, this is not a surprise.

There were also three other factors that were approaching significance in their correlation with the identification of /s/-aspiration, each with a *p*-value around .085. The first was a positive correlation between accuracy of the identification of /s/-aspiration and the percent use of Spanish when speaking. The second was a positive correlation between /s/-aspiration accuracy and the time spent speaking Spanish with NSs of English. The third was a negative correlation between /s/-aspiration accuracy and the amount of time spent watching media in English. The first correlation indicates that those who reported speaking a greater percentage of Spanish than English during the semester also showed more identification of /s/-aspiration as /s/. The second correlation shows that those who reported spending more time speaking Spanish with native English speaking peers also identified /s/-aspiration as a variant of /s/ more often. Finally, the third correlation indicates that those who watched more media in English during the semester were less accurate at identifying /s/-aspiration as a variant of /s/.

The factors that did not correlate significantly with Time 2 accuracy in the aspirated condition were the reported percent use of Spanish while reading, time spent speaking with NSs of Spanish in English, time spent speaking with NSs of English in English, watching media in Spanish, and amount of

time spent reading in Spanish and English. The findings of the correlation analysis for the forced-choice identification task will be discussed together with those of the lexical decision task after the discussion of the findings of the lexical decision task.

Language Use and Lexical Access. The Time 2 accuracy and Time 2-Time 1 percent change in accuracy for the learners' performance in the Asp real word condition of the lexical decision task were tested for correlation against the language use factors. The results of the analysis showed that Time 2 accuracy in the Asp real word condition was significantly correlated with three factors. The strongest was a negative correlation with the amount of time spent watching media in English. The second strongest correlation was a negative correlation with the amount of reported time spent speaking English with native speakers of Spanish. The final significant correlation was a positive correlation with the amount of time spent reading in Spanish. When considering the percent change over time in the Asp real word condition, there was a significant positive correlation with the reported time spent speaking English with native English speakers and no other significant correlations.

First, the negative correlation between Time 2 accuracy in the Asp real word condition and time spent watching media in English shows that the study abroad learners that reported spending more time on an average day watching media in English also exhibited lower accuracy of lexical access for words containing /s/-aspiration. Second, the negative correlation between reported time spent speaking English with native speakers of Spanish and Time 2 accuracy tells us that those who spent more time speaking English with NSs of Spanish exhibited lower accuracy for aspirated words. Finally, the positive correlation with time spent reading in Spanish indicates that those who spent more time reading in Spanish also had higher accuracy rates when responding to Asp real words. Additionally, two other significant negative correlations between the factors of watching media in English and reading in Spanish and speaking Spanish with native English speakers showed that those who reported more time

watching English-speaking media were also reading less in Spanish and speaking less in Spanish with English-speaking peers.

Factors that were not significantly correlated with Time 2 accuracy for the lexical decision task were the percent use of Spanish overall and while writing, reading, and speaking; the time spent speaking Spanish with NSs of Spanish, time spent speaking Spanish with NSs of English, time spent speaking English with NSs of English, watching media in Spanish, and reading in English. The implications of the results of the language use factors for both analyses will now be discussed together.

Summary of Target Language Use Factors for Both Tasks

When comparing the correlations between language use factors according to each task's accuracy, it is notable that there was only one factor that was correlated with accuracy in the aspirated condition for both tasks: Time spent watching media in English. This was a negative correlation for both tasks. However, the correlation was statistically significant only for the lexical decision task and was approaching significance for the forced-choice identification task. This consistent pattern of negative correlation for both perception tasks indicates that using media in English had an overall negative relationship with being able to correctly identify /s/-aspiration and access words that contained /s/-aspiration. One potential explanation is that more exposure to the L1 during SA, even via the media, may hinder L2 perceptual processing in the L2 in some way. Fox and McGory (2007), for example, suggested that media exposure to Northern American English may have contributed to their finding that Japanese speakers' perception of Northern American English vowels was superior to their perception of Southern American English vowels despite having lived in Alabama for two years. While this is not a direct comparison to study abroad learners using media in the L1, it does suggest that media use may negatively affect L2 perception when the language or dialect of the media is different from the target language or dialect that the individual is exposed to elsewhere. After all, there is evidence in the research to date that lexical access involves the lexicons of all of the languages that the speaker knows

in parallel (see Van Hell & Tanner, 2012 for a review), and the implication could be that exposure to the L1 could have an impact on L2 lexical access.

However, it also may not be the use of English-speaking media directly that affects perception. In the case of study abroad learners, it may be related to other factors such as individual language learning characteristics and goal-setting (Allen, 2010). It could be that L2 learners that spend more time watching media in English spend less time on activities that will provide them with L2 input, and that this could be related to their reasons and goals for studying abroad. Interestingly, there were significant negative correlations between the use of media in English and the reported time spent speaking in Spanish with NSs of Spanish, reading in Spanish and speaking in Spanish with NSs of English, which indicates that this could be a possible explanation. If those who spend more time watching media in English are also reading less in Spanish and speaking Spanish less with NSs of Spanish and English-speaking peers, this may indicate that these learners either did not prioritize L2 learning while in the study abroad context and had non-language-related reasons for studying abroad. It could, though, also signal other individual differences in personality or event context-related issues of homesickness or degree of acculturation, either of which could affect the activities in which the learners engaged and consequently their language acquisition.

The next finding of the correlation analysis of language use factors was that only one task, the forced-choice identification task, showed positive correlations between accuracy in the aspirated condition and the percent use of Spanish overall and the factors related to speaking Spanish. First of all, some of these factors were correlated with one another, showing that those who reported more Spanish use overall also reported more time spent speaking Spanish with NSs of Spanish, as would be expected. Next, it is important to think about why these factors were only significant for the forced-choice identification task and not the lexical decision task. A potential explanation is that the Asp condition of the forced-choice identification task exhibited a much lower overall percentage of accuracy

in tandem with a higher degree of variability among the study abroad learners than the Asp real word condition of the lexical decision task. This likely resulted from the use of only non-words in the design of the forced-choice identification task, which in combination with the presence a sociolinguistic variant was likely somewhat 'unnatural' compared to the lexical decision task, given that all other conditions of the forced-choice identification task demonstrated very high accuracy that was higher than the accuracy in the control conditions of the lexical decision task. Thus, the forced-choice identification task was an easier task in general, but not when considering /s/-aspiration. This is probably because non-words lack the phonological and lexical representations that are used to process lexical items. This means that the correlation analyses for both tasks were accounting for the same extralinguistic factors, such as speaking Spanish with NSs of Spanish, but that because the lexical decision data showed higher overall accuracy and less variability in the aspirated condition than the forced-choice identification task, the slopes of the correlations were flatter than those of the forced-choice identification task, diminishing the explanatory power of some of the extralinguistic factors. Thus, it seems that the difference in the significance of the extralinguistic factors according to task lies in the inherent differences between the tasks.

The difference between the two tasks in the significance of the extralinguistic factors is important for the interpretation of the data because it shows that overall, taking the results of both tasks together, most of the extralinguistic factors were not significant predictors of the perception of /s/-aspiration because the findings were not consistent across tasks. If both tasks showed similar accuracy rates and similar degrees of variability for aspirated items as compared to other non-aspirated items, the relationship of extralinguistic factors would be the same and it would be possible to conclude that factors such as speaking Spanish with NSs or using more Spanish overall were positively related to the identification and lexical processing of /s/-aspiration. However, the discrepancy between the tasks adds complexity. The question, then, is which task can be considered the most representative? I argue that the lexical decision task is the most representative of the participants' actual perception of /s/-

aspiration because of the SEVILLE group's patterns on this task. In other words, NSs of the target dialect should, in general, be equally as accurate at perceiving /s/-aspiration as they are at perceiving the standard sibilant variant because their phonological and lexical representations would include both the sibilant and /s/-aspiration as allophonic realizations of the phoneme /s/. There may be some difference, as previous research has reported that non-standard variants can be somewhat more difficult to perceive than standard variants, even for NSs that produce the non-standard variant (Sumner & Samuel, 2005, 2009). The SEVILLE group's accuracy in the Asp and CodaS conditions of the forced-choice identification task, though, was not equal by a fairly large margin (approximately 25%), but accuracy in these two conditions was much more similar on the lexical decision task (i.e., 6% difference), which was not a statistically significant difference. Also, the fact that the Asp real word condition of the lexical decision task consisted of real words that have phonological and lexical representations rather than non-words, which do not, makes the L2 participants' performance in this condition more representative of what they encounter in real interactions with NSs.

If we consider the lexical decision task a more representative task, then we can interpret the results of the language use factors as indicating that using English-speaking media and speaking English with NSs of Spanish had a negative relationship with the lexical processing of words containing /s/-aspiration, and that reading more in Spanish was positively related with more accurate responses to those words. But the language use factors related to speaking the target language with NSs and NNSs did not have significant effects on the processing of lexical items containing /s/-aspiration. As was discussed previously, it is very possible that reading more in Spanish, like writing, is a part of a broader profile of certain L2 learners that engage more in different language-related activities with the goal of learning the L2, compared to learners with different characteristics and/or different goals for studying abroad that are not related to language learning. Furthermore, there is evidence that reading in the L2 leads to incidental vocabulary learning (Paribakht & Wesche, 1999), which could positively affect

learners' performance on a measure of lexical decision. Likewise, more use of English, whether through media or otherwise, could be part of a broader profile of learners with non-language-related reasons for studying abroad or the influence of other factors that have nothing to do with the learners' goals and activities. For example, even learners that study abroad with the goal of learning the L2 may have shy personalities or suffer from homesickness, and therefore may not engage in the same ways in the L2 as learners who are more social and have accommodated more easily to the L2 cultural context. For these reasons it is impossible to say that certain learners read or wrote more in the L2 and engaged less with English for the sole reason that they had language-learning goals while the others did not. The interaction of these factors in the study abroad context is complex.

Next, it is important to consider why the language use factors related to *speaking* the target language, specifically, did not exert a stronger and more consistent effect on the learners' perception of /s/-aspiration. One may expect that speaking the target language more with NSs also leads to more input that may contain the target phonological variant. However, this is not necessarily true. The first hypothesis regarding the lack of effect overall (i.e., that the tasks showed largely different patterns related to target language use) and for lexical access results is that, since the goal of the study was not sociolinguistic in nature, but rather focused on learners' encoding of the linguistic meaning of /s/-aspiration, the language contact and use questionnaire was not designed to elicit the minute details of the types of conversations the learners had, the social characteristics of the interlocutors, and the characteristics of the input that they received. Segalowitz et al. (2004) suggested that taking into account the *nature* of interactions in the target language during study abroad may be the key to really understanding how target language use and contact with NSs during study abroad affect language acquisition. Hence, the current study supports the assertion of Segalowitz et al. (2004) that a general number of NS contacts and self-reported data on the percent use of the target language and time spent interacting with NSs in the target language are not robust enough to uncover consistent differences

among learners in terms of language acquisition, in this case the acquisition of an optional phonological variant. In the case of the acquisition of an optional variant this must be due, at least in part, to the nature of optional variants: that they are not categorically present in the input in every social context or for every NS of the target language.

According to exemplar theory, a variant that is not categorical will have a different cognitive representation than a normative variant because of the frequency with which they are present in the input of the L2 learners (Bybee, 2003). Applied to the current study, a normative variant such as the sibilant [s] will have a stronger phonological representation because of its frequency, while aspirated variants will be represented to different degrees for each learner depending on the input that they receive and the strength of the exemplar cluster. Understanding this, it should not be surprising that operationalizing language contact and use factors in ways that do not account for the individual characteristics of the interlocutors with whom the learners have contact and the contexts in which learners engage with NSs of the target language does not result in highly significant correlations between these factors and the learners' perception of a sociolinguistic variant. This is arguably going to be the most difficult methodological aspect of future research on the acquisition of sociolinguistic variants in the study abroad context. It not only necessitates more creative and robust methods of uncovering the details of L2 learners' interactions with NSs and then figuring out how to operationalize that information effectively as variables, but also requires detailed sociolinguistic research of the target variant in order for L2 researchers to better predict how the interlocutors with whom the learners come into contact will use the target variant.

The lack of a consistent effect overall for the language contact and use factors can also be explained in part by the nature of self-reported data. Self-reports do not necessarily coincide with what actually occurred, and may even be far from the truth in some cases. Other methods, though, such as first-hand observation of participants in the study abroad context and other such methods require much

more time, travel, and funding than most researchers can afford to invest. The ideal situation would be one in which the researcher can work with the L2 learners directly in the study abroad context in some capacity (e.g., study abroad program staff) other than being just 'the researcher' for the entire duration of the stay abroad, developing relationships with them and making detailed observations that can then be operationalized as variables in a study that has both quantitative and qualitative aspects. Clearly, though, this is very far from the realm of possibility for most researchers, representing the difficult balance in research on context of learning between methodological rigor and logistical reality. It will therefore likely take creative thinking in future research to come up with methods that can balance both of these in a more helpful way. This will be discussed further along with other limitations of the current study.

Another possible explanation for the lack of consistent language use effects is similar to what was discussed previously in relation to the similarity between the results of the participants in the two study abroad programs. In other words, the lack of consistent correlations across tasks for the language use factors, and the lack of significant correlations between factors related to speaking with NSs of Spanish and lexical access, might also be at least partially explained by an overall similarity among the L2 learners in terms of educational background, target language background, age, sex, and general L2 learning conditions in the study abroad context. This is not to say that there were no differences. However, the learners of both study abroad groups reported similar backgrounds related to the levels of Spanish language courses they had taken prior to SA, they were all of the same age group (i.e., university students), they were mostly women, and all came from similar university environments in the United States. As a group, then, the study abroad learners were quite homogenous. Furthermore, given the similar structure of the two study abroad programs, many of the learners likely had similar experiences during the semester of study abroad in terms of the types of activities in which they engaged and the types of people with whom they had contact. This is an inherent methodological issue when dealing

with populations of university students as research subjects. It clearly has the advantage of accessibility, since study abroad is a very common experience for college students in the United States and study abroad programs are plentiful. But it also suffers from a lack of both generalizability to other L2 learner populations and representing a wide enough range of individual learner characteristics. More differences based on language use factors and stronger correlations might be found when comparing learners in study abroad programs that are very distinct, or when comparing traditional study abroad students to other groups such as expatriates that are living and working in the same location as the study abroad students.

Summary of the Effects of Extralinguistic factors on Study Abroad Learners' Perception of /s/-Aspiration in Andalusian Spanish

The results of the correlation analyses testing whether study abroad program, grammar test score, language contact, and language use factors predicted gains over time showed that, in general, these factors did not exert a significant influence on how L2 learners in the study abroad context identified word-internal /s/-aspiration and accessed lexical items containing it. There were no significant differences between the learners based on study abroad program, since both groups' accuracy rates were very similar at the start of the semester abroad and both showed a very similar degree of positive change over time in the aspirated conditions of both perception tasks. This was hypothesized to have resulted from a high degree of similarity between the two study abroad programs and the learners in the two programs. However, it was demonstrated that while most of the learners in each study abroad group increased accuracy over time on both perception tasks, some also showed no change and others' accuracy in responding to aspirated stimuli decreased. To try to isolate factors that may have led to different individual results, the learners' grammatical proficiency, contact with NSs, and target language use were tested for correlation with accuracy at Time 2 and the percent change over time of accuracy.

In terms of grammatical proficiency, as measured by the grammar test, there was a significant correlation between the grammar test score at Time 1 and the learners' accuracy on the lexical decision task for all conditions except for the aspirated condition. This was also true of some conditions of the forced-choice identification task. The different effect for the aspirated condition compared to all other conditions demonstrated that learners with higher grammatical proficiency were able to perform at a higher level than those with lower grammatical proficiency in general, but the playing field was more level when words contained /s/-aspiration, not showing any advantage for having a higher grammar test score. Taking the results of the entire semester into account, there was no significant effect of grammatical proficiency over time on how the study abroad learners perceived /s/-aspiration, as the same pattern obtained at Time 2. Many learners showed some degree of increase in accuracy regardless of grammar test performance, showing that receiving input that contained /s/-aspiration was more crucial to their development than their proficiency.

In search of other factors to explain individual differences over time, the language contact and use factors were tested for correlation with accuracy in the aspirated condition for each task. The results showed different factors that were significant for each task. This was hypothesized to be due in part to inherent differences between the tasks, principally the differences in what the tasks required the participants to do and how closely that resembled the input and perceptual processing they experienced in real target language interactions. These differences were argued to be the reason for differences in accuracy and within-group variability in the aspirated condition of each task. Differences in accuracy and degree of variability in the aspirated conditions of the two tasks, then, were argued to have resulted in the differences between tasks of the significance of many of the language contact and use factors. Specifically, the fact that study abroad learners' mean accuracy in the aspirated condition of the lexical decision task was higher and showed less variability at both data collection times than the same condition of the forced-choice identification task meant that the slope of the lexical decision task

correlations was much flatter than for the forced-choice identification task, meaning that extralinguistic factors had less of a relationship to lexical decision accuracy than accuracy on the forced-choice identification task. It was argued, then, that the lexical decision task was more representative of what the learners would experience in real life in terms of input and perceptual processing, given that the task included real words with phonological and lexical representations that could change over time, while the forced-choice identification task did not.

For the forced-choice identification task, most of the positive correlations were related to using the target language more overall and spending more time speaking Spanish with NSs of Spanish. These factors, though, were not significantly correlated with lexical decision accuracy. For both tasks, either time spent reading in Spanish or writing in Spanish was significantly correlated with accuracy. These were hypothesized to be factors that might differentiate language learners with different characteristics and goals, which could affect the activities in which they engage during SA, which could affect the input they receive, and future research should continue to explore this. Another similarity between the tasks was that the time spent watching media in English was negatively correlated with accuracy in the aspirated condition to a significant or nearly-significant degree. This indicated that spending more time exposed to media in the L1 had a negative relationship with being able to identify and access words containing /s/-aspiration. This was hypothesized to have resulted from negative influences of the L1 on L2 perceptual processing and/or differences of language learning and individual characteristics among learners who expose themselves to more versus less L1 media during SA.

Finally, another important contribution of the analysis of extralinguistic factors, methodologically-speaking, was that the general inconsistency and lack of significant effects of extralinguistic factors that would be expected to have a significant effect is likely due to the design of the questionnaires and how learners' interactions with NSs of the target language are elicited and operationalized. Specifically, since optional (i.e., sociolinguistic) phonological variants are not

categorically present in the input in every social context and with every NS interlocutor, it will be necessary in future research to develop methods that will uncover individual differences in terms of the nature of the interactions between the L2 learners and NSs in the study abroad context and how much input that contains the target variant that L2 learners are actually getting.

Now that the research questions have been answered according to the findings of the current study, it is important to discuss the implications for our understanding of SLA during study abroad, as well as implications for the design of study abroad programs.

Implications of the Current Study for Our Understanding of SLA during Study Abroad

The literature reviewed in Chapters 1 and 2 showed that much of the research to date on SLA during study abroad has found that study abroad is a complex language learning context that involves many differences across individuals, study abroad programs, languages, and cultures. While many L2 learners believe they have made progress after a study abroad experience, the research to date has shown that the reason is not necessarily that they have better command of the L2 grammar or phonology after SA. Instead, as Segalowitz et al. (2004) argued, L2 learners most often ‘sound better’ because they have been shown to make substantial gains over time during study abroad in terms of oral fluency (e.g., Segalowitz & Freed, 2004; Valls-Ferrer, 2011; Valls-Ferrer & Mora, 2014), narrative discourse abilities (e.g., Collentine, 2004; Lafford, 2004), lexical complexity (e.g., Vives Boix & Meara, 2000; Juan-Garau, Salazar-Noguera & Prieto-Arranz, 2014) and communication strategies (e.g., Duperron & Overstreet, 2009; Lafford, 2004). In other words, the greatest benefit of the study abroad context, according to the research to date, is that it provides learners with opportunities to practice speaking and automatize their linguistic knowledge (DeKeyser, 2010) to some extent that allows for faster, more uninterrupted, and more complex-sounding speech. The intricacies of grammar and phonology, though, seem to be less of a priority to learners, whether consciously or subconsciously.

However, a recent body of research on the acquisition of variable forms during study abroad has brought about new questions and has produced evidence that acquiring variable forms is different from acquiring invariant aspects of the target language, and that this has application to language acquisition in the study abroad context.

Why might acquiring variation be different than acquiring invariant aspects of the target language during SA? The primary reasons are that sociolinguistic variants carry social meaning and distinguish some individuals and groups from others, while invariant linguistic forms do not, and that the use of sociolinguistic variants among NSs only occurs in certain dialect regions, among certain groups of people, and in certain social contexts. Research on sociolinguistic variation among speakers of the same L1 has shown that sociolinguistic variables can be categorized as social indicators, markers, and stereotypes that can be above or below the consciousness of the speakers of the language or dialect, depending on the type of variable (Labov, 1972). A sociolinguistic variable, depending on which type it is, will vary according to social and/or stylistic factors, and will also demonstrate variation according to linguistic factors such as phonological context, grammatical category of the word, and syllable stress. This can affect the strength of phonological representations according to Exemplar Theory (Bybee, 2003) and thus L2 learners' acquisition of a variant. Bayley and Regan (2004) argued that, "Variationist sociolinguistics...has suggested, convincingly in our view, that far from being a peripheral element, knowledge of variation is part of speaker competence. The implication of this position is that, in order to become fully proficient in the target language, second language learners also need to acquire native-speaker (NS) patterns of variation..." (p. 325). With the goal of investigating whether or not L2 learners acquire sociolinguistic variation in the second language and how it occurs, recent research has found that L2 learners in the study abroad context can, in fact, become sensitive to sociolinguistic variants and the constraints that govern their use by NSs of the target language (Geeslin et al., 2010, 2012; Geeslin, Fafulas & Kanwit, 2013; Kanwit & Solon, 2013; Linford, Zahler & Whatley, 2013; Salgado-Robles, 2011,

2014; Salgado-Robles & Enrique Ibarra, 2012). This has been primarily shown for morphosyntactic variants and contrasts with findings of previous work that has shown that L2 learners in the study abroad context have generally not demonstrated an advantage over traditional classroom learners for the acquisition of linguistic structures that are typically emphasized in classroom instruction. Regarding the acquisition of phonology during SA, we know that phonological gains have generally been minimal in the study abroad context both for sociolinguistic variants and other phonological forms (e.g., Avello & Lara, 2014; Diaz-Campos, 2004, 2006; Mora, 2014), given that L2 learners in this context often do not *produce* phonological sociolinguistic variants despite being exposed to them (Geeslin & Gudmestad, 2008b; Knouse, 2012; Ringer-Hilfinger, 2012). Consequently, one of the gaps in our knowledge about study abroad has been whether being immersed in the target language community has any effect over time on the L2 learners' *perception* of phonological variants during study abroad in the first place and what factors predict outcomes. Schmidt (2011) was an important first step in showing in a cross-sectional study that advanced L2 learners were able to identify /s/-aspiration as a variant of /s/ and that having more target language experience and having previously studied abroad in a /s/-weakening region were related to more identification of /s/-aspiration as a variant of /s/. But until the current study, we had not seen change over time in the perception of a phonological variant in Spanish by L2 learners in the study abroad context.

The current study, then, has taken the next step and has contributed to our understanding of SLA during study abroad by showing longitudinally that L2 learners who studied abroad made significant gains in how they identified /s/-aspiration and accessed words in the mental lexicon that contained /s/-aspiration, a dialect-specific phonological variant of Western Andalusian Spanish. Furthermore, this study has shown that studying abroad provides the input necessary to be able to encode the variant in phonological and lexical representations, while staying at one's home institution did not provide such input and, in turn, did not lead to gains for the phonological and lexical encoding of /s/-aspiration.

Though production was not tested, this is a different finding than what has been shown by studies such as George (2014), Knouse (2012) and Ringer-Hilfinger (2012). They found low rates of production of a dialect-specific phonological variant among L2 learners in the study abroad context and showed that factors such as social networks were important to determining whether they produced the variant or not. However, the current speech perception study has shown a different outcome, namely that the majority of the study abroad learners made gains over time for the perceptual encoding and processing of a dialect-specific phonological variant. We now have more evidence that language learning in the study abroad context provides opportunities for L2 learners to expand at least one aspect of their perceptual sociolinguistic competence (i.e., their understanding of the *linguistic* meaning of the variant) and become more well-rounded users of the L2 (Canale & Swain, 1980), regardless of whether they eventually incorporate a variant into their production. Future research must determine whether learners also acquire the social and contextual constraints surrounding the target variant. Thus, the current study is in accord with previous work that has also found evidence of learners acquiring other types of sociolinguistic variants in the study abroad context (e.g., Geeslin, Fafulas & Kanwit, 2013; Kanwit & Solon, 2013; Regan, 1995; Regan, Howard & Lemee, 2009; Sax, 2003; Salgado-Robles, 2011, 2014; Salgado-Robles & Enrique-Ibarra, 2012).

Implications for Study Abroad Program Design

As with any research, it is also important to offer practical applications of the findings, in this case so that study abroad programs can better help learners to acquire sociolinguistic competence in a second language. We have seen that most of the learners in both study abroad programs made advances over time, but that they still had more room for improvement at the end of the semester. Some previous research has found that explicit instruction, when combined with immersion, can lead to better outcomes than just immersion or instruction alone (e.g., Lord, 2010). Mora (2014) also suggested

that explicit instruction may be necessary to overcome the 'ceiling' that he observed in his perception study of study abroad learners. It may be of benefit to learners in the study abroad context to receive explicit instruction regarding the variety of the target language to which they are exposed during their time abroad. This could include aspects of grammar, phonology, pragmatics, the lexicon, and perhaps other aspects of language, which are unique to that particular region and people. This would not only increase the L2 learners' understanding of the local culture, people, and language variety, but also would help them to navigate the complexities of the study abroad context, such as the array of different communicative situations in which they find themselves. There has been a recent call for explicit instruction during study abroad related to dialect-specific features of pragmatics, and even a model for explicit instruction during study abroad proposed (Shively, 2010). However, I argue that this is equally relevant to other aspects of language acquisition, such as phonetics/phonology and morphosyntax.

Another suggestion for study abroad program design is that study abroad programs may be of more benefit to L2 learners if they specialize in reaching certain types of students rather than trying to reach all types of students. It has been suggested through Allen's (2010) research and others' that students study abroad for a variety of reasons. Combining students in the same program or classes that have very different goals for studying abroad could be detrimental to second language acquisition. If there are students who want a cultural and travel experience while getting course credit, perhaps they should be placed into programs and/or classes with others who have the same goal. For students whose focus is more on second language acquisition, they would likely be better served by a program in which they learn alongside others with the same goal and are given instruction and opportunities that will help them achieve that goal. This is surely more easily said than done, but perhaps study abroad programs could take even small steps toward trying to help students with different goals reach those goals in different ways.

One final suggestion for study abroad program design is to continually seek new ways to get students engaged meaningfully with native speakers of the target language, spend less time with peers that speak the same L1, and spend less time using media in the L1. This is particularly important for those who have the goal of learning the L2. The current study found that for both of the speech perception tasks, in the aspirated condition, those who spent more time engaging with media in the L1 also exhibited lower accuracy rates. This is indicative of either a L1 perceptual interference issue or differences in individual learner characteristics. It is not clear which, given that the current study did not investigate further. But in either case, it certainly could not hurt for study abroad programs to think of creative ways to provide meaningful opportunities for target language input to students and reduce their interaction with the L1.

Limitations and Future Directions

The current study was the first longitudinal study of the perception of a dialect-specific phonological variant in Spanish in a study abroad context. The goal of this study was to lay a foundation for future research on the perception of such variants in order to better understand how the L2 phonological system processes variation and what factors predict perceptual learning. As with any research study, it is necessary to address the limitations of the research and how future studies can improve on the methodology. The following are limitations of the design and implementation of the current study that should be considered for the design of future studies.

Participants. One limitation regarding the participants in the current study is that the two Study Abroad groups and the At-home group were different in terms of their prior Spanish course experience and this was reflected in a difference of their grammar test scores at Time 1 and the fact that half of the participants of each study abroad group had three or more years of university-level Spanish, while the at-home learners had not had more than two years of college Spanish. Recruiting the at-home students

from second-year courses was done to control the prior study abroad experience of the At-home group. However, knowing now what study abroad learners reported in terms of their prior language experience, it would be better for future research to try to match that experience for an At-home group. This was difficult to do beforehand because it was not known at the time what types of language experience the study abroad learners would have had prior to study abroad since they came from different universities with different courses and backgrounds. This is always a difficulty when recruiting participants for any SLA study and trying to maintain equilibrium among the groups. However, with more studies such as the current study, we can begin to better predict the level at which study abroad learners tend to begin a semester abroad and to better match At-home groups.

Study Abroad Programs. Another limitation of the current study is that it compared two very similar study abroad programs, rather than different study abroad program types. This limits the generalizability of the results to learners enrolled in different types of study abroad programs that may be more focused on service learning or some other feature of program design that provides different opportunities for meaningful interaction with NSs.

Task Design. A few limitations of the current study have to do with task design. For one, as was mentioned previously, the creation of the non-words for the lexical decision task sometimes consisted of a change of vowel and sometimes a change of consonant. A more consistent method for the creation of non-word stimuli would minimize the chances of misperceptions due to some differences between real and non-words being more salient than others due to the type of change that was made to create the non-word. Such a method could consist of only changing consonants and only using consonants that are of a place of articulation that is at least a certain distance away from that of the consonant in the real word. For example, a real word with a word-initial bilabial consonant could derive a non-word with a word-initial velar consonant. This is not always as simple as it seems because this procedure would inevitably result in the creation of some real words. It would take some trial and error, but would

certainly increase the consistency of the relationships between real words and their non-word counterparts, which would minimize the chances of incorrect responses that were due to the non-word being too phonetically similar to its real word counterpart.

Another task design limitation, which may not actually be a limitation and would require further research to find out, is that all stimuli were recorded with a rising intonation in order to avoid devoicing at the end of the words. While this has the advantage of stimuli that do not fall off at the end of the recording, it also means that the stimuli that were heard by the participants were produced with intonation that was question-like. The implication is that the stimuli were only technically representative of how question words would be produced by NSs of the target dialect, and it is likely that the L2 learners experienced the target words in question form much less frequently than in declarative form. Whether this actually had an effect on the data is not clear, but is a consideration for future studies. Replications of the current study using declarative intonation would help us to understand if this had an effect on the results.

Also, the fact that the forced-choice identification task was an untimed identification task that allowed the participants to repeat each stimulus once can be considered a limitation of that task. There was ample time for the participants to choose their response, processing demands were low, and the use of metalinguistic knowledge would have therefore been possible. Schmidt (2011) used the same type of identification task and remarked in her discussions of its limitations that perhaps tasks with greater processing demands, such as lexical decision or priming tasks, would show larger differences between native aspirating groups and non-aspirating NS and L2 groups. The current study did implement a lexical decision task and, interestingly, the differences between the groups were actually smaller and accuracy overall was higher on the lexical decision task than performance on the forced-choice identification task in the aspirated condition. In the sibilant and other control conditions, though, the accuracy on the forced-choice identification task was generally higher. The higher accuracy for

aspirated words in the lexical decision task occurred despite the greater time pressure of the task, which increases processing demands. The real difference between the tasks, therefore, is not hypothesized to be due to differences in processing demands, but rather the use of all non-words in the forced-choice identification task and the inclusion of half real words in the lexical decision task in combination with differences in what each task required of the listener in order to respond. The forced-choice identification task required identifying non-words orthographically, which can be a strange task since non-words are not something people experience in daily life. This could be especially strange with the presence of a geographic and sociolinguistic variant. At the same time, the lexical decision task required only that the listener determine whether a stimulus was a real or non-word with the push of a button. In this way, the lexical decision task's requirement for non-words was seemingly more natural, targeting more rapid processing. Additionally, real words are easier to process than non-words, as demonstrated by generally faster RTs for real words, and it therefore makes sense that even the aspirated real words would exhibit higher accuracy rates in lexical decision than non-word orthographic identification. After all, real aspirated words do contain other phonemes apart from /s/ that can be used to determine the identity of the lexical item as words compete for recognition, whereas there is no such advantage in the identification of non-words when all of the available orthographic options only differ according to the coda consonant and there is no phonological or lexical representation for any of the non-words. It seems, then, that the forced-choice identification task has the greatest limitations of the two tasks.

The tasks in the current study were also limited to isolated words that lacked sentential and discourse context. Given that speech is not composed of only isolated words, this restricts the generalizability of the findings. The current study, then, does not tell us how learners access lexical items in running speech, but only in isolation. Other contextual cues present in running discourse would certainly help learners to access lexical items containing /s/-aspiration. Furthermore, the words chosen for the current study focused solely on /s/-aspiration in word-internal position and not across word

boundaries when preceding a word-initial voiceless stop, where /s/-aspiration also occurs in Western Andalusian Spanish. It cannot be said, then, that the current study's findings can be generalized to all word positions.

Finally, the types of information that the language contact and use questionnaire elicited was also a limitation of the design of the current study. Specifically, it elicited self-reported data. While self-reported data is used very frequently in research on context of learning, it has certain well-known limitations. Self-reported data does not necessarily reflect what actually occurred. For one, thinking back over an entire semester and calculating averages on the spot to represent language use during the semester is a difficult task. Another reason is that learners can purposefully misrepresent their language use during the semester. Self-reported data is also difficult because, even though the instructions of the questionnaire gave the participants specific guidelines as to how to answer the questions, some learners did not respond correctly, reporting much larger values for time spent using the target language on an average day than are realistic. An attempt was made to control for these differences by creating percentages of the total time rather than using the reported time in hours, but we must consider that the language contact and use factors were measured by self-reports when interpreting the data.

One potential method for avoid eliciting self-reported data could be to determine who each participant's closest friend is in their study abroad program, or roommates, and have them evaluate each other's target language use. Logistically, this would not be much more difficult than obtaining self-reported data, and it would likely help to minimize the negative effects of self-reports. This approach, though, does have drawbacks, such as differences among learners in terms of their social lives, which may make it more difficult to find a complete group of pairs of learners that know one another well enough to make dependable evaluations. There could also be a similar effect to that of self-reported data if some learners represent their friends in a more positive light than might be warranted. It is clearly not a perfect solution, but could be an avenue of exploration for future research.

The questionnaire also lacked sufficient detail regarding the nature of the interactions that learners had in the target language during their semester abroad and the characteristics of the interlocutors with whom they interacted (Segalowitz et al., 2004). Such detail is even more important to studying the acquisition of sociolinguistic variants than other types of linguistic forms, particularly when investigating how learners acquire the social and contextual constraints of a sociolinguistic variant, which was not the goal of the current study.

Future Directions. With these limitations in mind, there are important questions that future research can answer with regard to the acquisition of phonological variation during study abroad. First, future studies should investigate both the perception and production of a phonological variant over time in the study abroad context in order to see how they are related and whether there are differences in their development over time. Research to date, including the current study, indicates that they are different, since the current study has shown development over time for the perception of a phonological variant while previous studies have shown very little production of phonological variants (e.g., Geeslin & Gudmestad, 2008b; Knouse, 2012; Ringer-Hilfinger, 2012). Research that combines quantitative and qualitative methods and both perception and production would provide a more complete picture of how phonological variants are acquired during study abroad and what factors predict both their perception and production.

It will also be important to investigate L2 learners' acquisition of the social and contextual constraints surrounding sociolinguistic variants during study abroad and not just the acquisition of linguistic meaning. On the one hand, the current study has shown that study abroad learners do acquire the perception and classification of /s/-aspiration as a variant of the phoneme /s/ over time, and that this helps their lexical access when confronted with words that contain the variant. Thus, we know that they learned to understand one aspect of the meaning of /s/-aspiration during study abroad (i.e., the linguistic meaning). But it remains to be seen whether the learners understand the social values that NSs

of the target variety place on /s/-aspiration and the appropriateness of its use in different speech contexts. This is important for L2 learners' overall sociolinguistic competence, particularly in relation to the production of sociolinguistic variants. Learners may incorporate a variant into the phonemic and lexical representations of the L2 phonological system and then learn how to accurately produce, doing so in the appropriate linguistic contexts (e.g., word and syllable positions), but at the same time they may not completely understand the social meaning attached to the variant and the contextual constraints on its use. For example, they may not know whether the variant is associated with being of a certain socioeconomic class or education level or whether it is considered informal and inappropriate to use in certain speech contexts. As Geeslin (2011) remarked, L2 learners who do not understand the social meanings attached to a variant "may project an inappropriate social image in certain contexts" (Geeslin, 2011b: 462). The learners could, then, come across in a negative way to the NSs with whom they interact and this, in turn, could affect their social relations with NSs and potentially the input that they receive.

In order to research L2 learners' acquisition of the social values related to the target variant, it will be necessary to develop more robust methods to uncover the nature of learners' interactions with NSs, the characteristics of the NS interlocutors with whom the learners have contact during SA, and employ sociolinguistic methods such as matched guise to determine whether learners are able to acquire an understanding of the social values that NSs place on the target variant. It will also be necessary to investigate learners' productions of the variant in different social contexts. This could be accomplished through the use of role play tasks or even observations in different communication situations and with different interlocutors. A possible avenue for future research could be to compare interviews with learners in which the interviewer is a variable. This could involve having learners be interviewed by native speakers of different characteristics, such as professors compared to teenagers, in order to see how the productions of the L2 learners vary in each situation. In this case, it would be

important to compare the amount of /s/-aspiration that is present in the speech of each type of interlocutor during the interview and whether that predicts learners' /s/-aspiration. Eventually, it will be important to incorporate the three components (i.e., perception, production, and the social component) into one research design in order to see the total development of sociolinguistic competence over time in the study abroad context. One final and important aspect of future research on the acquisition of variants that include a social component is that different variants are salient to different degrees, since they can be at the level of sociolinguistic indicators, markers, or stereotypes. It will be important in future research to account for differences in the level of awareness of the variant as part of the methodological design of research studies.

Another interesting future research question is how exposure to one type of /s/-aspiration, such as the kind studied by Schmidt (2011), affects the perception of that type of /s/-aspiration compared to how it affects the perception of another type of /s/-aspiration that has different acoustic cues, such as Western Andalusian /s/-aspiration. This is similar to what Ruch and Harrington (2014) did by testing how NSs of Argentinian Spanish, who produce /s/-aspiration with voiceless frication preceding a voiceless stop and normal VOT duration for the voiceless stop, perceived Western Andalusian /s/-aspiration, which consists of different acoustic cues. This type of research would provide clues as to how different acoustic cues present in the input are processed by L2 learners and whether exposure to some cues helps learners to perceive all types of /s/-aspiration or just the variety to which they are exposed.

Future research should also investigate how L2 learners perceive /s/-aspiration in sentential and discourse contexts, as well as across word boundaries. The current study's focus on isolated word-internal /s/-aspiration is a first step, but in order to understand how sentential and discourse cues affect learners' processing of /s/-aspiration it is necessary to expand the target context. Previous research has found that native speakers often use contextual information outside of the noun phrase that contains

/s/-aspiration or deletion to process the meaning of words in which /s/ elided (e.g., Hochberg, 1986; Hundley, 1987; Poplack, 1980; Ranson, 1991). It is important, then, to do the same for L2 learners.

Along similar lines, future research should investigate how learners weight differently the acoustic cues in their processing of Western Andalusian /s/-aspiration by manipulating the duration of VOT, the duration of the stop closure, and the duration of the vowel preceding /s/ in order to understand which cues learners use to encode /s/-aspiration. For example, Ruch and Harrington (2014) manipulated the duration of aspiration preceding a voiceless stop, as well as VOT, in order to determine how Argentinian speakers of Spanish weighted these acoustic cues when exposed to Andalusian /s/-aspiration. The same type of study would be helpful for understanding L2 learners' perception of this type of /s/-aspiration and whether their cue weighting changes over time when exposed to /s/-aspiration during SA. This is important because it can help us to understand if learners and native speakers use the same acoustic cues to process /s/-aspiration and whether learners become more nativelike in their processing over time.

Finally, since past research has shown differences in how learners of different L2 proficiency levels process lexical items using different parts of the brain and different degrees of cognitive control, future research should investigate the cognitive processing of non-standard variants compared to standard variants using technologies such as ERP and fMRI, including comparisons of L2 learners before and after study abroad to determine how the brain processes sociolinguistic variants and whether the same types of cognitive changes over time that have been reported in previous studies (e.g., Stein et al., 2009) are also found for non-standard variants.

Conclusions

To conclude, the current study has made contributions to the fields of second language acquisition, psycholinguistics, and sociolinguistics. For one, the current study has advanced our

understanding second language acquisition, specifically the acquisition of dialect-specific phonological variants by L2 learners. This study has demonstrated that the presence of a phonological variant in the speech signal, in this case /s/-aspiration in Andalusian Spanish, hinders L2 learners' perceptual processing more than the standard sibilant variant prior to exposure and that phonologically encoding /s/-aspiration and making gains over time for the access of lexical items containing it is possible when L2 learners are exposed to it through immersion in the target language community during study abroad. This is the first study to show this longitudinally for the perception of a dialect-specific phonological variant in Spanish and is in accord with previous research on the acquisition of morphosyntactic variants during study abroad (Geeslin et al., 2010, 2012; Geeslin, Fafulas & Kanwit, 2013; Kanwit & Solon, 2013; Salgado-Robles, 2011, 2014; Salgado-Robles & Enrique Ibarra, 2012). This has implications for our understanding of how second language variation is acquired in the study abroad context because just being immersed in the target dialect was sufficient to lead to the majority of the L2 learners making perceptual gains over time. This is in contrast to previous research that has shown that just being in the target dialect is not sufficient for learners to produce phonological variants, highlighting that the perception and production of phonological variants by L2 learners who are exposed to them are different. The most obvious reason for this comes from findings of previous studies such as Geeslin and Gudmestad (2008b), George (2014), Knouse (2012) and Ringer-Hilfinger (2012) that showed that L2 learners tend to need a reason to produce a dialect-specific variant before they will do so. George (2014) showed that the makeup of learners' social networks, whether they were primarily composed of NSs from the target dialect or from another dialect, played an important role in determining whether a learner would produce the variant or not. The results of the current study demonstrate that the perception of a dialect-specific variant in the study abroad context occurs and we can conclude that studying abroad, and thus receiving input containing the target variant, enhances the perceptual aspect of sociolinguistic competence. This is important because it shows that learners' comprehension of the

speech that is directed to them improves over time, which is certainly a positive outcome of studying abroad. This likely also has positive implications for the L2 learners' cultural understanding and appreciation.

This study has also contributed to our understanding of study abroad because it was shown that L2 learners of similar backgrounds that were enrolled in comparable types of traditional study abroad programs showed equivalent patterns of acquisition over time. First, this is important because it shows that phonological gains are possible in a traditional three-month study abroad program. Previous research has suggested that a longer period of time may be necessary to see phonological gains in during study abroad with respect to the perception of difficult phonemic contrasts (i.e., non-optional contrasts) and production (e.g., Avello & Lara, 2014; Mora, 2014). Studies showed a general lack of significant gains over time in the traditional study abroad context, but that learners that stayed abroad for an extended period of time, such as Alvord and Christiansen's (2012) missionaries, demonstrated more gains. We have now seen that when acquiring the *perception* of an optional phonological variant, at least one that is as perceptually salient as /s/-aspiration, significant gains can be made in just under three months' time. Second, it is important because it shows that studying a relatively homogenous group of learners in the study abroad context leads to findings that are not significantly different among them. This has implications for the types of programs that study abroad research targets in future studies. It will be important to consider in future research the different types of study abroad programs and the advantages and disadvantages of each type of program for L2 learners' input and linguistic outcomes. Recent research has begun to compare program types (e.g., Dewey, Bown et al., 2014; Martinsen et al., 2010), but we need to continue to seek connections between study abroad program type to actual linguistic outcomes.

The final main implications of this study to our understanding of SLA in a study abroad context have to do with extralinguistic factors, what they tell us, and how they are measured. First, only one

extralinguistic factor that was strongly correlated with both the identification of /s/-aspiration and the lexical access of words containing /s/-aspiration (lexical decision task) was the amount of time the learners spent being exposed to media in the L1 during SA. This finding contributes to our understanding of language acquisition during study abroad by providing evidence of a negative effect of exposure to the L1 on how learners encode and process a dialect-specific phonological variant. It was argued that this effect could be related either to L1 perceptual interference, which would not be surprising given previous findings that both the L1 and L2 lexical items compete in the process of lexical access, or to learner characteristics that lead L2 learners with non-language-related goals or have other individual differences (e.g., personality) to spend more time engaging with media in the L1. Additionally, the fact that writing and reading in the target language during study abroad were significantly and positively correlated with accuracy for either the forced-choice identification task or lexical decision task and also positively correlated to some degree with speaking the target language with NSs suggests a connection between these activities and learner characteristics, such as differences in language proficiency (vocabulary size, grammatical proficiency) or goals for study abroad, which may have an effect on the perception of dialect-specific variants.

The analysis of extralinguistic factors also showed that the methods that have been commonly used to measure L2 learners' interactions with NSs and target language use, such as self-reported time spent engaging with NSs in the target language and the number of NS contacts a learner has, cannot adequately account for the acquisition of sociolinguistic variation. This is because sociolinguistic variants are not present in the input all of the time and in every communicative context. For this reason, new methods must be created that target the *underlying nature* of the interactions (Segalowitz et al., 2004) and predict the frequency with which the target variant is actually present in the input that each learner receives. Only then will it be possible to tease apart individual differences related to contact with NSs and target language use and have a greater likelihood of seeing significant effects on the results.

This study has also made an important contribution to L2 psycholinguistics because it demonstrated that, for one, the optional phonological variant was more difficult for L2 learners to process than the standard sibilant variant. This was found both for the accuracy of lexical decision and response time, indicating a high degree of processing cost for the optional variant. A second important finding was that the processing of the optional variant changed over time due to immersion. And a third finding is that L2 grammatical proficiency, as measured by a grammar test, was positively correlated to higher accuracy of lexical access for all conditions of the lexical decision task except for the aspirated condition. This shows that, while those of higher L2 proficiency have a general processing advantage in terms of being able to access words in the mental lexicon, they did not have an advantage for processing the optional variant. In other words, the playing field seems to be level for optional variants in terms of cognitive processing. This is important for L2 psycholinguistics research because it provides impetus for future research geared toward discovering how the underlying cognitive processing mechanisms work when processing optional variants compared to standard variants, and whether there is change over time in terms of the underlying cognitive mechanisms that L2 learners use to process optional variants when they are exposed to them.

The implications of this research for L2 psycholinguistics are also methodological. This dissertation has demonstrated that the use of multiple different speech perception tasks that measure different levels of phonological processing is crucial to understanding how a dialect-specific phonological variant is encoded and processed. This was shown because a task that has been very commonly used in L2 phonology research, forced-choice identification using orthographic representations and non-words, was shown to produce results that were similar in some ways to the lexical decision task and different in others. Importantly, it was found that the use of non-words in combination with orthographic representations seemed to have a negative effect on the results due to creating some confusion among the participants as to how to categorize /s/-aspiration with non-words. But the identification task did

reveal patterns of association that were helpful for interpreting the data. Given the drawbacks of orthographic identification, though, the inclusion of a lexical decision task was important because it was able to show what occurs when real words containing /s/-aspiration were experienced by the participants and how it affected their lexical access of words for which there were actual lexical and phonological representations. The importance of the use of multiple tasks is in accord with previous research on the acquisition of variable forms in a second language (e.g., Geeslin & Gudemestad, 2008a) and has again been shown to be useful in understanding how another aspect of linguistic competence, speech perception, develops over time.

Finally, this dissertation has implications for sociolinguistic research as well. First, it is the only study to date to test how native speakers of Andalusian Spanish encode and process /s/-aspiration compared to the standard sibilant variant. This is important because, though the research on /s/-aspiration in Spanish has covered a wide range of topics and dialects, including descriptions of /s/-aspiration in this Andalusian Spanish, this study has expanded our understanding of one of this distinctive feature of Andalusian Spanish, including how NSs of that dialect encode and process the phonological variation that is prevalent in that dialect. This study also, although it was not the main goal of the research questions, has contributed further evidence of perceptual differences between NSs of different dialects that differ in their use of the target sociolinguistic variant (e.g., Boomershine, 2006; Schmidt, 2011; Sumner & Samuel, 2009). Specifically, it was shown that non-aspirating NSs were less accurate than aspirating NSs only for the perception of /s/-aspiration, but not other control conditions. However, non-aspirating NSs were still able to access aspirated words in the mental lexicon a good percentage of the time, lending further support to findings that it is not necessary for a speaker-listener to incorporate dialect-specific variants into their production to be able to perceive them (Schmidt, 2011; Sumner & Samuel, 2009).

Through the findings and contributions of this dissertation it is hoped that researchers in the fields of second language acquisition, psycholinguistics, L2 phonology, sociolinguistics, and those working at the intersection of all of these fields, can develop new research that further advances our understanding of how sociolinguistic phonological variation is processed by native speakers and L2 learners, and the role of different contexts of learning in the acquisition of variation in a second language.

References

- Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20(3), 242-275.
- Adamson, H. D., & Regan, V. M. (1991). The acquisition of community speech norms by Asian immigrants learning English as a second language. *Studies in Second Language Acquisition*, 13(1), 1-22.
- Alba, O. (1982). Función del acento en el proceso de elisión de la /s/ en la República Dominicana. In O. Alba (Ed.), *El Español del Caribe: Ponencias del VI Simposio de Dialectología* (17–26). Santiago: Universidad Católica Madre y Maestra.
- Alba, O. (2000). Variación de la /s/ en las noticias de televisión. In O. Alba (Ed.), *Nuevos aspectos del español en Santo Domingo*, (27–48). Santo Domingo: Librería La Trinitaria.
- Alba, O. (2004). *Cómo Hablamos Los Dominicanos: Un Enfoque Sociolingüístico*. Dominican Republic: Grupo León Jimenes.
- Alfaraz, G. (2000). *Sound Change in a Regional Variety of Cuban Spanish*. (Unpublished doctoral dissertation). Michigan State University, East Lansing.
- Aliaga-García, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. In M. Watkins, S. Rauber, & B. O. Baptista (Eds.), *Recent research in second language phonetics/phonology: Perception and production* (2-31). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.
- Aliaga-García, C., Mora, J. C., & Cerviño-Povedano, E. (2011). L2 speech learning in adulthood and phonological short-term memory. *Poznań Studies in Contemporary Linguistics PSiCL*, 47(1), 1-14.

- Allen, H. W. (2010). Language-learning motivation during short-term study abroad: An activity theory perspective. *Foreign Language Annals*, 43(1), 27-49.
- Alvord, S. M., & Christiansen, D. E. (2012). Factors influencing the acquisition of Spanish voiced stop spirantization during an extended stay abroad. *Studies in Hispanic and Lusophone Linguistics*, 5(2), 239-276.
- Avello, P., & Lara, A. R. (2014) Phonological development in L2 speech production during study abroad programmes differing in length of stay. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts* (137-166). Amsterdam: John Benjamins.
- Baker, W., & Smith, L. C. (2010). The impact of L2 dialect on learning French vowels: Native English speakers learning Quebecois and European French. *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, 66(5), 711-738.
- Barron, A. (2006). Learning to say 'you' in German: The acquisition of sociolinguistic competence in a study abroad context. In M. A. DuFon & E. Churchill (Eds.), *Language learners in study abroad contexts*, (59-90). Tonawanda, NY: Multilingual Matters.
- Bataller, R. (2010). Making a request for a service in Spanish: Pragmatic development in the study abroad setting. *Foreign Language Annals*, 43(1), 160-175.
- Bayley, R. (2005). Second Language Acquisition and Sociolinguistic Variation. *Intercultural Communication Studies*, 14(2), 1-11.
- Bayley, R. & Regan, V. (2004). Introduction: The acquisition of sociolinguistic competence. *Journal of Sociolinguistics*, 8, 323-338.
- Bayley, R. & Tarone, E. (2012). Variationist approaches to second language acquisition. In P. Robinson (Ed.), *Routledge Encyclopedia of Second Language Acquisition*. New York: Routledge.

- Beattie, J., Valls-Ferrer, M., & Pérez-Vidal, C. (2014). Listening performance and onset level in formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts* (195-216). Amsterdam: John Benjamins.
- Beebe, L. M. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning, 30*(2), 433-445.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *Haskins Laboratories Status Report on Speech Research, 107/108*, 1-30.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research* (171-204). Baltimore: York Press.
- Best, C. T., Hallé, P. A., Bohn, O. S., & Faber, A. (2003). Cross-language perception of nonnative vowels: Phonological and phonetic effects of listeners' native languages. *Proceedings of the 15th International Congress of Phonetic Sciences, 2889-2892*. Barcelona, Spain.
- Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*(3), 305-330.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second language speech perception. Commonalities and complementarities. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (13-34). Philadelphia: John Benjamins.
- Bialystok, E., & Craik, F. I. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science, 19*(1), 19-23.
- Bishop, J. B. (2007). Incomplete neutralization in Eastern Andalusian Spanish: Perceptual consequences of durational differences involved in s-aspiration. In *Proceedings of the 16th International Congress of Phonetic Sciences, 1765-1768*. Saarbrücken, Germany.

- Boersma, P., & Weenink, D (2014). Praat: doing phonetics by computer [Computer program]. Version 5.3.77, retrieved 18 May 2014 from <http://www.praat.org/>.
- Boomershine, A. (2006). Perceiving and Processing Dialectal Variation in Spanish: An Exemplar Theory Approach. In T. L. Face & Klee, C. A. (Eds.), *Selected Proceedings of the 8th Hispanic Linguistics Symposium* (58-72). Somerville, MA: Cascadilla Proceedings Project.
- Bradley, T. (2006). Phonetic realizations of /sr/ clusters in Latin American Spanish. In M. Diaz-Campos (Ed.), *Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology* (1-13). Somerville, MA: Cascadilla Proceedings Project.
- Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, 61(5), 977-985.
- Brecht, R., Davidson, D., & Ginsberg, R. (1995). Predictors of foreign language gain during study abroad. In B. Freed (Ed.), *Second language acquisition in a study abroad context*, (37-66). Amsterdam: John Benjamins.
- Browman, C.P.; Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3-4), 155-180.
- Brown, E. L., & Cacoullos, R. T. (2002). ¿ Qué le vamos a hacer?: Taking the syllable out of Spanish /s/ reduction. *University of Pennsylvania Working Papers in Linguistics*, 8(3), 17-31.
- Brown, E. L., & Cacoullos, R. T. (2003). Spanish [s]. In R. Núñez-Cedeño & L. López (Eds.), *A Romance perspective on language knowledge and use: Selected papers from the 31st Linguistic Symposium on Romance Languages* (21-38). Amsterdam: John Benjamins.
- Brysbaert, M. (1998). Word recognition in bilinguals: Evidence against the existence of two separate lexicons. *Psychologica Belgica*, 38, 163-165.

- Bybee, J. (2000). Lexicalization of sound change and alternating environments. In M. Broe & J. Pierrehumbert (Eds). *Papers in Laboratory Phonology 5: Acquisition and the lexicon*, (250-268). Cambridge: Cambridge University Press.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14(03), 261-290.
- Bybee, J. (2003). *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, J. (2007). *Frequency of use and the organization of language*. New York: Oxford University Press.
- Campbell-Kibler, K. (2010). New directions in sociolinguistic cognition. *University of Pennsylvania Working Papers in Linguistics*, 15(2), 5.
- Calero, M. A., (1993). *Estudio sociolingüístico del habla de Toledo*. Lleida, España: Pagés.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Carbonero, P. (1982). *El habla de Sevilla*. Seville, Spain: Biblioteca de Temas Sevillanos.
- Carlet, A., & Cebrian, J. (2014). Training Catalan speakers to identify L2 consonants and vowels: A short-term high variability training study. *Concordia Working Papers in Applied Linguistics*, 5, 85-98.
- Carlson, K. M. (2012). An acoustic and perceptual analysis of compensatory processes in vowels preceding deleted post-nuclear /s/ in Andalusian Spanish. *Concentric: Studies in Linguistics*, 38(1), 39-67.
- Carvalho, A. M. (2006). Spanish (s) aspiration as a prestige marker on the Uruguayan-Brazilian border. *Spanish in Context*, 3(1), 85-114.
- Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, 34(3), 372-387.
- Cedergren, H. C. J. (1973). *The interplay of social and linguistic factors in Panama*. (Unpublished doctoral dissertation). Cornell University: Ithaca, NY.

- Cepeda, G. (1995). Retention and deletion of word-final /s/ in Valdivian Spanish (Chile). *Hispanic Linguistics*, 6/7, 329–353.
- Cheng, A. C., & Mojica-Diaz, C. C. (2006). The effects of formal instruction and study abroad on improving proficiency: The case of the Spanish subjunctive. *Applied Language Learning*, 16(1), 17.
- Cid-Hazard, S. M. (2003). Variación de estilo en relación a la variable fonológica /s/ en el español de Santiago de Chile. *Southwest Journal of Linguistics*, 22(2), 13-43.
- Clayton, I. (2010). On the natural history of pre-aspirated stops (Unpublished doctoral dissertation). University of North Carolina: Chapel Hill, NC.
- Cohen, A. D., Paige, R. M., Shively, R. L., Emert, H., & Hoff, J. (2005). Maximizing study abroad through language and culture strategies: Research on students, study abroad program professionals, and language instructors. Minneapolis: Center for Advanced Research on Language Acquisition, University of Minnesota.
- Cohen, A. D., & Shively, R. L. (2007). Acquisition of requests and apologies in spanish and french: Impact of study abroad and strategy-building intervention. *The Modern Language Journal*, 91(2), 189-212.
- Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26(2), 227-248.
- Collentine, J., & Freed, B. F. (2004). Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition*, 26(2), 153-356.
- Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, 32(2), 193-210.
- Corder, S. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 5, 161-169.

- Davidson, D. E. (2010). Study abroad: When, how long, and with what results? New data from the Russian front. *Foreign Language Annals*, 43(1), 6-26.
- Davies, Mark. (2002) Corpus del Español: 100 million words, 1200s-1900s. Available online at <http://www.corpusdelespanol.org>.
- Davies, M. (2006). *A frequency dictionary of Spanish: Core vocabulary for learners*. NY: Routledge.
- DeKeyser, R. (1991). Foreign language development during a semester abroad. *Foreign language acquisition research and the classroom* (104-119). Lexington, MA: D.C. Heath.
- DeKeyser, R. (2010). Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals*, 43(1), 80-92.
- Dewaele, J. M. (2002). Using sociostylistic variants in advanced French interlanguage: The case of nous/on. *Eurosla Yearbook*, 2(1), 205-226.
- Dewaele, J. M. (2004a). Retention or omission of the *ne* in advanced French interlanguage: The variable effect of extralinguistic factors. *Journal of Sociolinguistics*, 8(3), 433-450.
- Dewaele, J. M. (2004b). Vous or tu? Native and non-native speakers of French on a sociolinguistic tightrope. *International Review of Applied Linguistics in Language Teaching*, 42(4), 383-402.
- Dewey, D. P. (2004). A comparison of reading development by learners of Japanese in intensive domestic immersion and study abroad contexts. *Studies in Second Language Acquisition*, 26(02), 303-327.
- Dewey, D. P., Bown, J., Baker, W., Martinsen, R. A., Gold, C., & Eggett, D. (2014). Language use in six study abroad programs: An exploratory analysis of possible predictors. *Language Learning*, 64(1), 36-71.
- Dewey, D. P., Ring, S., Gardner, D., & Belnap, R. K. (2013). Social network formation and development during study abroad in the Middle East. *System*, 41(2), 269-282.

- Diao, W., Freed, B., & Smith, L. (2011). Confirmed *beliefs or false assumptions?* A study of home stay experiences in the French study abroad context. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 21, 109-142.
- Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, 26(2), 249-274.
- Díaz-Campos, M. (2006). The effect of style in second language phonology: An analysis of segmental acquisition in study abroad and regular-classroom students. In T. L. Face and C. A. Klee (Eds.), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, 26-39. Somerville, MA: Cascadilla Proceedings Project.
- Dickerson, L. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, 9, 401–407.
- Dijkstra, T., & Van Heuven, W. J. (1998). The BIA model and bilingual word recognition. *Localist connectionist approaches to human cognition*, 189-225.
- Dijkstra, T., Van Jaarsveld, H., & Brinke, S. T. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, 1(01), 51-66.
- Di Silvio, F., Donovan, A., & Malone, M. E. (2014). The Effect of Study Abroad Homestay Placements: Participant Perspectives and Oral Proficiency Gains. *Foreign Language Annals*, 47(1), 168-188.
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing*. New York: Routledge.
- Du, H. (2013). The development of Chinese fluency during study abroad in China. *The Modern Language Journal*, 97(1), 131-143.
- Duperron, L., & Overstreet, M. H. (2009). Preparedness for study abroad: Comparing the linguistic outcomes of a short-term Spanish program by third, fourth and sixth semester L2 learners. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 18, 157-179.

- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1568-1578.
- Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A "deafening" deafness" in French?. *Journal of Memory and Language*, 36, 406-421.
- Dwyer, M. M. (2004). More Is Better: The Impact of Study Abroad Program Duration. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 151-163.
- Eilers, R. E., Bull, D. H., Oller, D. K., & Lewis, D. C. (1984). The discrimination of vowel duration by infants. *The Journal of the Acoustical Society of America*, 75(4), 1213-1218.
- Elliott, A. R. (1995). Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *The Modern Language Journal*, 79(3), 356-371.
- Ellis, R. (1985). Sources of variability in interlanguage. *Applied linguistics*, 6(2), 118-131.
- Engle, L., & Engle, J. (2003). Study abroad levels: Toward a classification of program types. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 9(1), 1-20.
- Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26(04), 551-585.
- Escudero, P. (2005). *Linguistic perception and second-language acquisition: Explaining the attainment of optimal phonological categorization* (Unpublished doctoral dissertation). Utrecht University, Utrecht, Netherlands.
- Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech*, 53(3), 343-365.
- Escudero, P., & Williams, D. (2011). Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. *The Journal of the Acoustical Society of America*, 129(1), EL1-EL7.

- Escudero, P., & Williams, D. (2012). Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *The Journal of the Acoustical Society of America*, 131(5), EL406-EL412.
- Fafulas, S. (2010). The acquisition of a morphosyntactic variable in Spanish: The analytic and synthetic verb forms of present progressive aspect. *Indiana University Linguistics Club Working Papers Online*, 9, 1-19.
- Fafulas, S. (2012). New perspectives on the variation in the simple present and present progressive tense in Spanish and English. *Spanish In Context*, 9(1), 58-87.
- Fafulas, S. (2013). *First and second-language patterns of variation: Acquisition and use of simple present and present progressive forms in Spanish and English* (Unpublished doctoral dissertation). Indiana University, Bloomington, IN.
- Félix-Brasdefer, J. C., & Koike, D. A. (Eds.). (2012). *Pragmatic variation in first and second language contexts: Methodological issues* (Vol. 31). Amsterdam: John Benjamins.
- Félix-Brasdefer, J. C., & Hasler-Barker, M. (2015). Complimenting in Spanish in a short-term study abroad context. *System*, 48, 75-85.
- Figuroa, N. (2000). An acoustic and perceptual study of vowels preceding deleted post-nuclear /s/ in Puerto Rican Spanish. In H. Campos, E. Herburger, A. Morales-Front, and T. J. Walsh (Eds.), *Hispanic Linguistics at the Turn of the Millennium: Papers from the 3rd Hispanic Linguistics Symposium*, (66-79). Somerville, MA: Cascadilla Press.
- Firth, A., & Wagner, J. (2007). On Discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal*, 91(1 [Supplement]), 757-772.
- Flege, J. E. (1984). The detection of French accent by American listeners. *The Journal of the Acoustical Society of America*, 76(3), 692-707.

- Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America*, 89(1), 395-411.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, (233-277). Baltimore: York Press.
- Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437-470.
- Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6(3), 185-202.
- Flege, J. E., & Hillenbrand, J. (1986). Differential use of temporal cues to the /s/-/z/contrast by native and non-native speakers of English. *The Journal of the Acoustical Society of America*, 79(2), 508-517.
- Flege, J. E., MacKay, I. R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, 106(5), 2973-2987.
- Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97(5), 3125-3134.
- Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of phonetics*, 29(2), 191-215.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41(1), 78-104.
- Foltz, D. (1991). A study of the effectiveness of studying Spanish overseas. *Paper presented at the Pennsylvania State Modern Language Association Annual Meeting*, Pittsburgh, PA.

- Fontanella de Weinberg, M. B. (1990). Variación y cambio lingüístico en el español bonaerense, II: Visión diacrónica (1580-1980). In B. Garza Cuaron (Ed.), *Homenaje a Jorge A. Suarez: Lingüística Indoamericana e Hispánica*. Mexico City: Centro de Estudios Lingüísticos y Literarios, Colegio de México.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116-124.
- Fox, R. A., & McGory, J. T. (2007). Second language acquisition of a regional dialect of American English by native Japanese speakers. *Language experience in second language speech learning: In honor of James Emil Flege*, 117-134.
- Fraser, C. C., (2002). Study abroad: an attempt to measure the gains. *German As a Foreign Language Journal*, 1, 45-65.
- Freed, B. (1990). Language learning in a study abroad context: The effects of interactive and non-interactive out-of-class contact on grammatical achievement and oral proficiency. In J. E. Alatis (Ed.), *Linguistics, language teaching and language acquisition: The interdependence of theory, practice and research*, (459-477). Washington, D. C.: Georgetown University Press.
- Freed, B. F. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context*, (123-148). Philadelphia: John Benjamins.
- Freed, B. F. (1998). An Overview of Issues and Research in Language Learning in a Study Abroad Setting. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 4(2), 31-60.
- Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition*, 26(02), 349-356.

- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26(02), 275-301.
- Freed, B., So, S., & Lazar, N. A. (2003). Language learning abroad: How do gains in written fluency compare with gains in oral fluency in French as a second language?. *ADFL BULLETIN*, 34(3), 34-40.
- Galarza, I., Delgado-Díaz, G., & Willis, E. (2014, April). *¿Nuevamente /s/? Una nueva mirada a la elisión de /s/ implosiva en el español de Puerto Rico*. Paper presented at the 7th Workshop on Spanish Sociolinguistics, Madison, WI.
- García, A. (2013). *Allophonic variation in the Spanish sibilant fricative* (Unpublished doctoral dissertation). University of Wisconsin Milwaukee, Milwaukee, WI.
- García-Amaya, L. (2009). New findings on fluency measures across three different learning contexts. In J. Collentine, M. García, B. Lafford & F. Marcos Marín (Eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, (68-80). Somerville, MA, Cascadilla Proceedings Project.
- García-Amaya, L. J. (2012). *Second language fluency and cognition: The study of Spanish second language development in an overseas immersion program and an at-home foreign language classroom* (Unpublished doctoral dissertation). Indiana University, Bloomington.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1993). Match and mismatch in phonological context. *Proceedings of the Fifteenth Annual Meeting of the Cognitive Science Society*, (446-451). Hillsdale, NJ: Erlbaum.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, 12(5-6), 613-656.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1999). Ambiguity, competition, and blending in spoken word recognition. *Cognitive Science*, 23(4), 439-462.

- Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, 45(2), 220-266.
- Gatbonton, E., & Trofimovich, P. (2008). The ethnic group affiliation and L2 proficiency link: Empirical evidence. *Language Awareness*, 17(3), 229-248.
- Geeslin, K. L. (2000). A new approach to the second language acquisition of copula choice in Spanish. In R. P. Leow & C. Sanz (Eds.), *Spanish Applied Linguistics at the Turn of the Millennium: Papers from the 1999 Conference on the L1 & L2 Acquisition of Spanish and Portuguese*, (50-66). Somerville, MA: Cascadilla Press.
- Geeslin, K. (2011a). The acquisition of variation in second language Spanish: How to identify and catch a moving target. In M. Díaz-Campos (Ed.), *Handbook of Spanish Sociolinguistics*, (303-320). Malden, MA: Wiley-Blackwell.
- Geeslin, K. L. (2011b). Variation in L2 Spanish: The state of the discipline. *Studies in Hispanic & Lusophone Linguistics*, 4(2), 461-517.
- Geeslin, K. L., & Fafulas, S. (2012). Variation of the simple present and present progressive forms: A comparison of native and non-native speakers. In K. Geeslin & M. Diaz-Campos (Eds.), *Selected Proceedings from the 14th Hispanic Linguistics Symposium*, (179-196). Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, K. L., Fafulas, S., & Kanwit, M. (2013). Acquiring geographically-variable norms of use: The case of the present perfect in Mexico and Spain. In C. Howe, S. E. Blackwell, & M. Lubbers Quesada (Eds.), *Selected Proceedings of the 15th Hispanic Linguistics Symposium*, (205-220). Somerville, MA: Cascadilla Proceedings Project.

- Geeslin, K. L., García-Amaya, L., Hasler-Barker, M., Henriksen, N., & Killam, J. (2010). The SLA of direct object pronouns in a study abroad immersion environment where use is variable. In C. Borgonovo, M. Español-Echevarría, & P. Prévost (Eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium* (246-259). Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, K. L., García-Amaya, L. J., Hasler-Barker, M., Henriksen, N. C., & Killam, J. (2012). The L2 acquisition of variable perfective past time reference in Spanish in an overseas immersion setting. In K. Geeslin & M. Diaz-Campos (Eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, (197-213). Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, K. L., & Gudmestad, A. (2008a). Comparing interview and written elicitation tasks in native and non-native data: Do speakers do what we think they do?. In J. Bruhn de Garavito & E. Valenzuela (Eds.), *Selected Proceedings of the 10th Hispanic Linguistics Symposium*, (64-77). Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, K. L., & Gudmestad, A. (2008b). The acquisition of variation in second-language Spanish: An agenda for integrating studies of the L2 sound system. *Journal of Applied Linguistics*, 5(2), 137-157.
- Geeslin, K. L., & Gudmestad, A. (2010). An exploration of the range and frequency of occurrence of forms in potentially variable structures in second-language Spanish. *Studies in Second Language Acquisition*, 32(03), 433-463.
- Geeslin, K. & Gudmestad, A. (2011). Using sociolinguistic analyses of discourse-level features to expand research on L2 variation: Native and non-native contrasts in forms of Spanish subject expression. In L. Plonsky and M. Schierloh (Eds.), *Selected Proceedings of the 2009 Second Language Research Forum: Diverse Contributions to SLA*, (16-30). Somerville, MA: Cascadilla Proceedings Project.

- Geeslin, K., Linford, B., Fafulas, S., Long, A., & Campos, M. D. (2013). The L2 development of subject form variation in Spanish: The individual vs. the group. In J. Cabrelli Amaro, G. Lord, A. de Prada Pérez, & J. E. Aaron (Eds.) *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, (156-174), Somerville, MA: Cascadilla Proceedings Project.
- George, A. (2014). Study abroad in central Spain: The development of regional phonological features. *Foreign Language Annals*, 47(1), 97-114.
- Giegerich, H. J. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press.
- Giles, H., Coupland, J., & Coupland, N. (Eds.). (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press.
- Gill, M., & Wegmann, B. (2014). *En contacto, enhanced student text: Gramática en acción*. Stamford, CT: Cengage Learning.
- Godfrey, L., Treacy, C., & Tarone, E. (2014). Change in French second language writing in study abroad and domestic contexts. *Foreign Language Annals*, 47(1), 48-65.
- Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia*, 9(3), 317-323.
- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, 29(2), 228-244.
- Gregg, K. R. (1990). The variable competence model of second language acquisition, and why it isn't. *Applied linguistics*, 11(4), 364-383.
- Gudmestad, A. (2006). L2 variation and the Spanish subjunctive: Linguistic features predicting mood selection. In C. A. Klee & T. L. Face (Eds.), *Selected proceedings from the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, 170-184. Somerville, MA: Cascadilla Proceedings Project.

- Gudmestad, A. (2008). *Acquiring a variable structure: An interlanguage analysis of second-language mood use in Spanish*. (Unpublished doctoral dissertation). Indiana University, Bloomington.
- Gudmestad, A. (2010a). Moving beyond a sentence-level analysis in the study of variable mood use in Spanish. *Southwest Journal of Linguistics*, 29(1), 25-51.
- Gudmestad, A. (2010b). The second-language acquisition of temporal distinctions within the subjunctive mood in Spanish. *Paper presented at the International Conference on Applied Linguistics*, Chiayi, Taiwan.
- Gudmestad, A. (2012a). 'Toward an understanding of the relationship between mood use and form regularity: Evidence of variation across tasks, lexical items, and participant groups. In K. Geeslin & M. Diaz-Campos (Eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, (214-227). Somerville, MA: Cascadilla Proceedings Project.
- Gudmestad, A. (2012b). Acquiring a variable structure: An interlanguage analysis of second language mood use in Spanish. *Language Learning*, 62(2), 373-402.
- Gudmestad, A., & Geeslin, K. L. (2010). Exploring the roles of redundancy and ambiguity in variable subject expression: A comparison of native and non-native speakers. In C. Borgonovo, M. Español-Echevarría, & P. Prévost (Eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium* (270-283). Somerville, MA: Cascadilla Proceedings Project.
- Gudmestad, A. & Geeslin, K. L. (2011). Assessing the use of multiple forms in variable contexts: The relationship between linguistic factors and future time reference in Spanish. *Studies in Hispanic and Lusophone Linguistics*, 4, 3-33.
- Gudmestad, A. & K. Geeslin. (2013). Second-language development of variable future-time expression in Spanish. In A. Carvahlo & S. Beaudrie (Eds.), *Refereed Proceedings of the 6th International Workshop on Spanish Sociolinguistics*, (63-75). Somerville, MA; Cascadilla Proceedings Project.

- Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, *107*(5), 2711-2724.
- Hallé, P. A., Best, C. T., & Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics*, *27*(3), 281-306.
- Hamann, S., Boersma, P., & Cavar, M. (2010). Language-specific differences in the weighting of perceptual cues for labiodentals. In K. Dziubalska-Kolaczyk, M. Wrembel & M. Kul (Eds.), *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, New Sounds 2010*, (167-172).
- Hammond, R. (1980). Las realizaciones fonéticas del fonema /s/ en el español cubano rápido de Miami. In G. E. Scavnick (Ed.), *Dialectología hispanoamericana: Estudios actuales*, (8-15). Washington D.C.: Georgetown University Press.
- Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *The Journal of the Acoustical Society of America*, *110*(1), 489-503.
- Hassall, T. (2013). Pragmatic development in study-abroad contexts. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*, *8*, (4516-4522). Hoboken, NJ: Wiley Blackwell.
- Henriksen, N. C., Geeslin, K. L., & Willis, E. W. (2010). The development of L2 Spanish intonation during a study abroad immersion program in León, Spain: Global contours and final boundary movements. *Studies in Hispanic and Lusophone Linguistics*, *3*(1), 113-162.
- Henríquez Ureña, P. (1921). Observaciones sobre el español de América. *Revista de Filología Española*, *8*, 357-390.
- Hernandez, T. A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *The Modern Language Journal*, *94*(4), 600-617.

- Hochberg, J. G. (1986). Functional compensation for/s/deletion in Puerto Rican Spanish. *Language*, 62, 609-621.
- Howard, M. (2006). Variation in advanced French interlanguage: A comparison of three (socio)linguistic variables. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 62(3), 379-400.
- Howard, M. (2011). Input perspectives on the role of learning context in Second Language Acquisition. An introduction to the special issue. *International Review of Applied Linguistics in Language Teaching*, 49(2), 71-82.
- Howard, M. (2012). The advanced learner's sociolinguistic profile: On issues of individual differences, second language exposure conditions, and type of sociolinguistic variable. *The Modern Language Journal*, 96(1), 20-33.
- Howard, M., Lemée, I., & Regan, V. (2006). The L2 acquisition of a phonological variable: The case of /l/ deletion in French. *Journal of French Language Studies*, 16(01), 1-24.
- Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge: Cambridge University Press.
- Hundley, J. E. (1987). Functional constraints on plural marker deletion in Peruvian Spanish. *Hispania*, 70(4), 891-894.
- Ife, A., Vives Boix, G., & Meara, P. (2000). The impact of study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics*, 4(1), 55-84.
- Ingram, J. C., & Park, S. G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, 25(3), 343-370.
- Institute of International Education. (2013). "Profile of U.S. Study Abroad Students, 2001/02-2011/12." Open Doors Report on International Educational Exchange. Retrieved from <http://www.iie.org/opendoors>.

- Isabelli, C. (2004a). *A case study of the factors in the development of Spanish linguistic accuracy and oral communication skills: Motivation and extended interaction in the study abroad context*. Lewiston, NY: Edwin Mellen Press.
- Isabelli, C. (2004b). Study abroad for advanced foreign language majors: Optimal duration for developing complex structures. In H. Byrnes & H. Maxim (Eds.), *Advanced foreign language learning: A challenge to college programs* (114-130). Boston: Heinle.
- Isabelli, C., & Nishida, C. (2005). Development of the Spanish subjunctive in a nine-month study-abroad setting. In D. Eddington (Ed.), *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, (78-91). Somerville, MA: Cascadilla Proceedings Project.
- Isabelli-García, C. (2010). Acquisition of Spanish gender agreement in two learning contexts: Study abroad and at home. *Foreign Language Annals*, 43(2), 289-303.
- Jiménez Sabater, M. (1975). *Más datos sobre el español de la República Dominicana*. Santo Domingo: Ediciones Intec.
- Juan-Garau, M., Salazar-Noguera, J., & Prieto-Arranz, J. I. (2014). English L2 learners' lexico-grammatical and motivational development at home and abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts* (235-258). Amsterdam: John Benjamins.
- Kanwit, M., & Geeslin, K. (2014). 'The interpretation of Spanish subjunctive and indicative forms in adverbial clauses: A Cross-sectional study. *Studies in Second Language Acquisition*, 36(3), 1-47.
- Kanwit, M., & Solon, M. (2013). Acquiring variation in future-time expression abroad in Valencia, Spain and Mérida, Mexico. In J. Cabrelli Amari, G. Lord, A. de Prada Perez, & J. E. Aaron (Eds), *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, (206-221). Somerville, MA: Cascadilla Proceedings Project.

- Kennedy, K. M. (2012). *What we don't learn in the classroom: The acquisition of sociolinguistic competence during study abroad*. (Unpublished doctoral dissertation). University of California, Davis, CA.
- Killam, J. (2011). *An interlanguage analysis of differential object marking in L2 Spanish* (Unpublished doctoral dissertation), Indiana University, Bloomington, IN.
- Kinginger, C. (2008). Language learning in study abroad: Case studies of Americans in France. *The Modern Language Journal*, 92(s1), 1-124.
- Kinginger, C., & Farrell, K. (2004). Assessing development of meta-pragmatic awareness in study abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 19-42.
- Kiparsky, P. (1972). Explanation in phonology. *Goals of Linguistic Theory*, 189(227), 454-68.
- Knight, S. M., & Schmidt-Rinehart, B. C. (2010). Exploring conditions to enhance student/host family interaction abroad. *Foreign Language Annals*, 43(1), 64-71.
- Knouse, S. M. (2012). The acquisition of dialectal phonemes in a study abroad context: The case of the Castilian theta. *Foreign Language Annals*, 45(4), 512-542.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Kuhl, P. (1992). Psychoacoustics and speech perception: Internal standards, perceptual anchors, and prototypes. In Lynne A. Werner & Edwin W. Rubel (Eds.), *Developmental Psychoacoustics*, (293-332). Washington, D.C.: American Psychological Association.
- Kuhl, P. (1993). Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics*, 21, 125-139.
- Kuhl, P. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, 4, 812-822.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, 97(22), 11850-11857.

- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.
- Lafford, B. (1986). Valor diagnóstico-social del uso de ciertas variantes de /s/ en el Español de Cartagena, Colombia. In H. Cedergren, R. A. Núñez Cedeño, I. Páez Urdaneta and J. M. Guitart (Eds.), *Estudios Sobre la Fonología del Español del Caribe*, (53-74). Caracas: Ediciones La Casa de Bello.
- Lafford, B. A. (1995). Getting into, through and out of a survival situation: A comparison of communicative strategies used by students studying Spanish abroad and 'at home'. In B. F. Freed & C. A. Ferguson (Eds.), *Studies in Bilingualism*, 9, 97-121. Amsterdam: John Benjamins.
- Lafford, B. A. (2004). The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition*, 26(02), 201-225.
- Lafford, B. A. (2006). The effects of study abroad vs. classroom contexts on Spanish SLA: Old assumptions, new insights and future research directions. In C. A. Klee & T. L. Face (Eds.), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (1-25). Somerville, MA: Cascadilla Proceedings Project.
- Lafford, B., & Collentine, J. (2006). The effects of study abroad and classroom contexts on the acquisition of Spanish as a second language. In R. Salaberry & B. Lafford (Eds.), *The Art Of Teaching Spanish: Second Language Acquisition: From Research To Praxis* (103-126). Washington D.C: Georgetown University Press.
- Lavoie, L. M. (2000). *Phonological patterns and phonetic manifestations of consonant weakening*. (Unpublished doctoral dissertation). Cornell University, Ithaca, NY.
- Lenneberg, E. H. (1969). On explaining language. *Science*, 164, 635-643.

- Lewis, M. Paul, Simons, G.F., & Fennig, C.D. (eds.). (2014). *Ethnologue: Languages of the World*, Seventeenth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- Linford, B. G. (2009). *The acquisition of subject pronouns in second language Spanish* (Unpublished Masters thesis). University of Montana, Missoula.
- Linford, B. G. (In preparation). *The second-language development of dialect-specific morpho-syntactic variation in Spanish during study abroad*. (Unpublished doctoral dissertation). Indiana University, Bloomington, IN
- Linford, B & Shin, S. (2013). Lexical frequency effects on L2 Spanish subject pronoun expression. In J. Cabrelli Amaro, G. Lord, A. de Prada Pérez, & J. E. Aaron (Eds.), *Selected Proceedings of the of the 16th Hispanic Linguistics Symposium*, (175-189). Somerville, MA: Cascadilla Proceedings Project.
- Linford, B. G., Zahler, S., & Whatley, M. (October, 2013). The impact of study abroad on L2 Spanish null vs. overt subject pronoun variation. *Paper presented at the 32nd Second Language Research Forum*, Provo, UT.
- Lipski, J. M. (1984). On the weakening of /s/ in Latin American Spanish. *Zeitschrift für Dialektologie und Linguistik*, 31-43.
- Lipski, J. M. (1985). /s/ in Central American Spanish. *Hispania*, 68, 143-149.
- Lipski, J. (1989). /s/-voicing in Ecuadoran Spanish: Patterns and principles of consonantal modification. *Lingua*, 79(1), 49-71.
- Lipski, J. (1995). Literary 'Africanized' Spanish as a research tool: Dating consonant reduction. *Romance Philology*, 49:2, 130-167.
- Lipski, J. (1999). The many faces of Spanish /s/-weakening: (re) alignment and ambisyllabicity. In J. Gutiérrez-Rexach & J. Martínez Gil (Eds.), *Advances in Hispanic Linguistics*, 198-213.

- Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and speech, 10*(1), 1-28.
- Liskin-Gasparro, J. (1984). Comparison of the oral proficiency of students of Spanish in a study abroad program and those in regular academic programs. As cited in Liskin-Gasparro and Beyer (1987).
- Liskin-Gasparro, J., & Beyer, T. (1987). The effect of intensive-immersion conditions of the acquisition and development of oral proficiency in Russian and Spanish. *Grant Proposal to the Department of Education*.
- Llanes, À., & Muñoz, C. (2009). A short stay abroad: Does it make a difference?. *System, 37*(3), 353-365.
- Llanes, À., & Muñoz, C. (2013). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning, 63*(1), 63-90.
- Long, M. H. (1981). Input, interaction, and second language acquisition. *Annals of the New York Academy of Sciences, 379*(1), 259-278.
- Long, M. H. (1997). Construct validity in SLA research: A response to Firth and Wagner. *The Modern Language Journal, 81*(3), 318-323.
- Long, M. (1998). SLA: Breaking the siege. *University of Hawai'i working papers in ESL, 17*(1), 79-129.
- Lopez Ortega, N. (2003). *The development of discourse competence in study abroad learners: A study of subject expression in Spanish as a second language*. (Unpublished doctoral dissertation). Cornell University, Ithaca.
- Lord, G. (2009). Second-language awareness and development during study abroad: A case study. *Hispania, 92*(1), 127-141.
- Lord, G. (2010). The Combined Effects of Immersion and Instruction on Second Language Pronunciation. *Foreign Language Annals, 43*(3), 488-503.
- Luce, P. A., & McLennan, C. T. (2005). Spoken word recognition: The challenge of variation. In D. Pisoni & R. Remez (Eds.), *The Handbook of Speech Perception*, (590-609). Hoboken, NJ: Wiley Blackwell.

- Luna, K. (2010). *The Spanish of Ponce, Puerto Rico: A phonetic, phonological, and intonational analysis*. (Unpublished Doctoral Dissertation). University of California, Los Angeles.
- Lyster, R., & Rebuffot, J. (2002). Acquisition des pronoms d'allocution en classe de français immersif. *Acquisition et interaction en langue étrangère, 17*, 51-72.
- Ma, R., & Herasimchuk, E. (1975). The linguistic dimensions of a bilingual neighborhood. In J. Fishman (Ed.), *Bilingualism in the Barrio*, (347-479). Bloomington, IN: Indiana University Press.
- MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics, 2*(04), 369-390.
- Magnan, S. S. (1986). Assessing speaking proficiency in the undergraduate curriculum: Data from French. *Foreign Language Annals, 19*(5), 429-438.
- Magnan, S. S., & Back, M. (2007). Social interaction and linguistic gain during study abroad. *Foreign Language Annals, 40*(1), 43-61.
- Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. New York: Routledge.
- Malovrh, P. A. (2006). L2 sentence processing of Spanish OVS word order and direct object pronouns: An analysis of contextual constraints. In N. Sagarra & A. J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, (169-179). Somerville, MA: Cascadilla Proceedings Project.
- Malovrh, P. A. (2008). *A multifaceted analysis of the interlanguage development of Spanish direct-object clitic pronouns observed in L2-learner production* (Unpublished doctoral dissertation), Indiana University, Bloomington.
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition, 6*(02), 97-115.
- Marqués-Pascual, L. (2011). Study abroad, previous language experience, and Spanish L2 development. *Foreign Language Annals, 44*(3), 565-582.

- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1), 71-102.
- Martinsen, R. A. (2010). Short-term study abroad: Predicting changes in oral skills. *Foreign Language Annals*, 43(3), 504-530.
- Martinsen, R. A. (October, 2013). *Language Learning in an Immersion Environment*. Paper presented at the Second Language Research Forum, Provo, UT.
- Martinsen, R. A., & Alvord, S. M. (2012). On the relationship between L2 pronunciation and culture. *Spanish in Context*, 9(3), 443-465.
- Martinsen, R. A., Alvord, S. M., & Tanner, J. (2014). Perceived Foreign Accent: Extended Stays Abroad, Level of Instruction, and Motivation. *Foreign Language Annals*, 47(1), 66-78.
- Martinsen, R. A., Baker, W., Bown, J., & Johnson, C. (2011). The benefits of living in foreign language housing: The effect of language use and second-language type on oral proficiency gains. *The Modern Language Journal*, 95(2), 274-290.
- Martinsen, R. A., Baker, W., Dewey, D. P., Bown, J., & Johnson, C. (2010). Exploring diverse settings for language acquisition and use: Comparing study abroad, service learning abroad, and foreign language housing. *Applied Language Learning*, 20(1-2), 45-69.
- Mason, K. W. (1994). *Comerse las eses: A Selective Bibliographic Survey of /s/ Aspiration and Deletion in Dialects of Spanish* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor.
- Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition*, 13(3), 279-297.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive psychology*, 18(1), 1-86.
- McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(4), 539.

- Möhle, D. (1984). A comparison of the second language speech of different native speakers. In H. Dechert et al. (Eds.), *Second Language Productions*, (26-49). Tübingen: Gunter Narr.
- Mora, J. C. (2014). The role of onset level on L2 perceptual phonological development after formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*, (167-194). Amsterdam: John Benjamins.
- Mora, J. C., & Valls-Ferrer, M. (2012). Oral Fluency, Accuracy, and Complexity in Formal Instruction and Study Abroad Learning Contexts. *TESOL Quarterly*, 46(4), 610-641.
- Morris, R. E. (2000). Constraint interaction in Spanish /s/-aspiration: Three peninsular varieties. In H. Campos, E. Herburger, A. Morales-Front & T. Walsh (Eds.), *Hispanic Linguistics at the Turn of the Millennium: Papers from the 3rd Hispanic Linguistics Symposium*, (14-30). Somerville, MA: Cascadilla Proceedings Project.
- Mougeon, R., Nadasdi, T. & Rehner, K. (2010). *The sociolinguistic competence of immersion students*. Bristol, UK: Multilingual Matters.
- Mougeon, R., Rehner, K., and Nadasdi, T. (2004). The learning of spoken French variation by immersion students in Toronto, Canada. *Journal of Sociolinguistics*, 8, 408–432.
- Motohashi-Saigo, M., & Hardison, D. M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Technology*, 13(2), 29-47.
- Moya Corral, J. A. (2007). Noticia de un sonido emergente: la africada dental procedente del grupo *st* en Andalucía. *Revista de filología de la Universidad de La Laguna*, 25, 457-466.
- Muñoz, C., & Llanes, À. (2014). Study abroad and changes in degree of foreign accent in children and adults. *The Modern Language Journal*, 98(1), 432-449.
- Navarro Tomás, T. (1918). *Manual de pronunciación española*. Madrid: Consejo Superior de Investigaciones Científicas.

- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29(04), 557-581.
- O'Connor, N. (1988). Oral proficiency testing of junior year abroad: Implications for the undergraduate curriculum. *Paper presented at the 1988 Annual Meeting of the Modern Language Association*.
- O'Neill, P. (2009). S-aspiration and occlusives in Andalusian Spanish: Phonetics or phonology? In O. Parker Jones & E. Payne (Eds.), *Oxford University Working Papers in Linguistics, Philology & Phonetics*, 12, 73-85.
- O'Neill, P. (2010). Variación y cambio en las consonantes oclusivas del español de Andalucía. *Estudios de fonética experimental*, 19, 11-41.
- Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. *Studies in Second Language Acquisition*, 21(02), 195-224.
- Parrell, B. (2012). The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of Phonetics*, 40(1), 37-45.
- Pérez-Vidal, C., & Barquin, E. (2014). Comparing progress in academic writing after formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*, (217-234). Amsterdam: John Benjamins.
- Pérez-Vidal, C., & Juan-Garau, M. (2009). The effect of study abroad (SA) on written performance. *Eurosla Yearbook*, 9(1), 269-295.
- Pérez-Vidal, C., & Juan-Garau, M. (2011). The effect of context and input conditions on oral and written development: A study abroad perspective. *International Review of Applied Linguistics in Language Teaching*, 49(2), 157-185.

- Pérez-Vidal, C., Juan-Garau, M. and Mora, J.C. (2011). The effects of formal instruction and study abroad contexts on foreign language development: the SALA project. In C. Sanz, C. & R. Leow (Eds.), *Implicit and Explicit Conditions, Processes, and Knowledge in SLA and Bilingualism*, (194-215). Georgetown: Georgetown University Press.
- Petrides, M. (1998). Specialized systems for the processing of mnemonic information within the primate frontal cortex. In A. C. Roberts, T. W. Robbins, & L. Weiskrantz (Eds.), *The prefrontal cortex—executive and cognitive functions* (103–116). Oxford, UK: Oxford University Press.
- Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. L. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure*, (137-157). Amsterdam: John Benjamins.
- Pisoni, D. B., & Levi, S. V. (2007). Some observations on representations and representational specificity in speech perception and spoken word recognition. In G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics*, 3-18. Oxford, UK: Oxford University Press.
- Pisoni, D. B., & Lively, S. E. (1995). Variability and invariance in speech perception: A new look at some old problems in perceptual learning. In W. Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research* (429-455). Baltimore: York Press.
- Pizziconi, B. (2013). Japanese vocabulary development in and beyond study abroad: the timing of the year abroad in a language degree curriculum. *The Language Learning Journal*, 41, 1-20.
- Polivanov, E. D. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle linguistique de Prague*, 4, 7-96.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421-435.
- Poplack, S. (1980). Deletion and disambiguation in Puerto Rican Spanish. *Language*, 56(2), 371-385.

- Poplack, S. (1986). Acondicionamiento gramatical de la variación fonológica en un dialecto puertorriqueño. In H. Cedergren, R. A. Núñez Cedeño, I. Páez Urdaneta, & J. M. Guitart (Eds.), *Estudios Sobre la Fonología del Español del Caribe* (95-107). Caracas: Ediciones La Casa de Bello.
- Quilis, A. (1999). *Tratado de fonética y fonología españolas*. Gredos.
- Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 57(2), 273-298.
- Ranson, D. L. (1991). Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change*, 3(2), 133-152.
- Raupach, M. (1984). Formulae in second language speech production. In H. Dechert, D. Möhle & M. Raupach (Eds.), *Second Language Production*, (26-49). Tübingen: Gunter Narr.
- Regan, V. (1995). The acquisition of sociolinguistic native speech norms: Effects of a year abroad on second language learners of French. In B. F. Freed & C. A. Ferguson (Eds.), *Studies in Bilingualism*, (245-267). Amsterdam: John Benjamins.
- Regan, V. (1996). *Variation in French interlanguage: A longitudinal study of sociolinguistic competence*. Amsterdam: John Benjamins.
- Regan, V. (2003). Sociolinguistics and language learning in a study abroad context. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 4, 61-90.
- Regan, V., Howard, M., & Lemée, I. (2009). *The acquisition of sociolinguistic competence in a study abroad context* (Vol. 40). Bristol, UK: Multilingual Matters.
- Rehner, K. (2002). The development of aspects of linguistic and discourse competence by advanced second language learners of French. (Unpublished doctoral dissertation). OISE/University of Toronto, Toronto, Canada.

- Rehner, K., & Mougeon, R. (1999). Variation in the spoken French of immersion students: To *ne* or not to *ne*, that is the sociolinguistic question. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 56(1), 124-154.
- Rehner, K., Mougeon, R. & Nadasdi, T. (2003). The learning of sociolinguistic variation by advanced FSL learners. *Studies in Second Language Acquisition*, 25(1), 127-156.
- Rickford, J. R., & McNair-Knox, F. (1994). Addressee-and topic-influenced style shift: A quantitative sociolinguistic study. In D. Biber (Ed.), *Sociolinguistic Perspectives on Register*, (235-276).
- Ringer-Hilfinger, K. (2012). Learner acquisition of dialect variation in a study abroad context: The case of the Spanish [θ]. *Foreign Language Annals*, 45(3), 430-446.
- Rivers, W. P. (1998). Is being there enough? The effects of homestay placements on language gain during study abroad. *Foreign Language Annals*, 31(4), 492-500.
- Rose, M. (2012). Cross-Language Identification of Spanish Consonants in English. *Foreign Language Annals*, 45(3), 415-429
- Rosner, B. S., López-Bascuas, L. E., García-Albea, J. E., & Fahey, R. P. (2000). Voice-onset times for Castilian Spanish initial stops. *Journal of Phonetics*, 28(2), 217-224.
- Ruch, H. (2012). Affrication of /st/: Western Andalusian Spanish: variation and change from a sociophonetic point of view. In S. Calamai, C. Celata and L. Ciucci (Eds.), *Proceedings of Sociophonetics, at the Crossroads of Speech Variation, Processing and Communication*, (61-64).
- Ruch, H. (2013). Investigating a gradual metathesis: Phonetic and lexical factors on/s/-aspiration in Andalusian Spanish. *University of Pennsylvania Working Papers in Linguistics*, 19(2), 171-180.
- Ruch, H., & Harrington, J. (2014). Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics*, 45, 12-25.
- Ruiz-Sánchez, C. (2004). El comportamiento de la /s/ implosiva en el habla de Caracas. *Boletín de Lingüística*, 21, 48-65.

- Ruiz-Sánchez, C. (2005). Restricciones funcionales en la elisión del marcador de plural en el español de Caracas. *Boletín de Lingüística*, 23, 89-105.
- Salgado-Robles, F. (2011). *The acquisition of sociolinguistic variation by learners of Spanish in a study abroad context* (Unpublished doctoral dissertation), University of Florida, Gainesville.
- Salgado-Robles, F. (2014). Variación dialectal por aprendientes de español en un contexto de inmersión en el extranjero: Un análisis cuantitativo del uso léxico en el discurso oral y escrito. *Lenguas Modernas*, 43, 97-112.
- Salgado-Robles, F., & Enrique Ibarra, C. (2012). "Les voy a echar de menos cuando regrese a los Estados Unidos": Adquisición de la variación dialectal por aprendientes de español en un contexto de inmersión. *Ogigia*, 11, 61-77.
- Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning*, 54(3), 525-582.
- Sasaki, M. (2007). Effects of study-abroad experiences on EFL writers: A multiple-data analysis. *The Modern Language Journal*, 91(4), 602-620.
- Sasaki, M. (2009). Changes in English as a foreign language students' writing over 3.5 years: A sociocognitive account. *Writing in foreign language contexts: Learning, teaching, and research*, 49-76.
- Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' L2 writing ability and motivation: A Longitudinal Study. *Tesol Quarterly*, 45(1), 81-105.
- Sax, K. J. (2003). *Acquisition of stylistic variation in American learners of French* (Unpublished doctoral dissertation). Indiana University, Bloomington.
- Schauer, G. (2007). Finding the right words in the study abroad context: The development of German learners' use of external modifiers in English. *Intercultural Pragmatics*, 4(2), 193-220.

- Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26(6), 1270-1281.
- Schmidt, L. (2009). The effect of dialect familiarity via a study abroad experience on L2 comprehension of Spanish. In J. Collentine, M. García, B. Lafford & F. Marcos Marín (Eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, 143-154. Somerville, MA: Cascadilla Proceedings Project.
- Schmidt, L. B. (2011). *Acquisition of dialectal variation in a second language: L2 perception of aspiration of Spanish /s/* (Unpublished doctoral dissertation). Indiana University, Bloomington.
- Schmidt, L. B., & Willis, E. W. (2011). Systematic investigation of voicing assimilation of Spanish /s/ in Mexico City. In S. Alvord (Ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, 1-20.
- Schmidt-Rinehart, B. C., & Knight, S. M. (2004). The homestay component of study abroad: Three perspectives. *Foreign Language Annals*, 37(2), 254-262.
- Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, 22(1), 106-129.
- Scott, D. R., & Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behavior*, 23(4), 450-466.
- Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26(02), 173-199.
- Segalowitz, N., Freed, B., Collentine, J., Lafford, B., Lazar, N., & Díaz-Campos, M. (2004). A comparison of Spanish second language acquisition in two different learning contexts: Study abroad and the domestic classroom. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 1-18.

- Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (371–388). Oxford, UK: Oxford University Press.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209-231.
- Serrano, R., Llanes, A., & Tragant, E. (2011). Analyzing the effect of context of second language learning: domestic intensive and semi-intensive courses vs. study abroad in Europe. *System*, 39(2), 133-143.
- Serrano, R., Tragant, E., & Llanes, A. (2012). A longitudinal analysis of the effects of one year abroad. *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, 68(2), 138-163.
- Sessarego, S. (2011). Phonetic analysis of /sr/ clusters in Cochabambino Spanish. In L. A. Ortiz (Ed.), *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, 251-263. Somerville, MA: Cascadilla Proceedings Project.
- Shively, R. L. (2011). L2 pragmatic development in study abroad: A longitudinal study of Spanish service encounters. *Journal of Pragmatics*, 43(6), 1818-1835.
- Shively, R. L. (2013). Out-of-class interaction during study abroad: Service encounters in Spain. *Spanish in Context*, 10(1), 53-91.
- Shoemaker, E. M. (2010). Nativelike attainment in L2 listening: The segmentation of spoken French. In M. Dziubalska-Kolaczyk, M. Wrembel, & M. Kul (Eds.), *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech. New Sounds 2010*, Poznan. Poznan: Adama Mickiewicz University.
- Silverman, D. (2003). On the rarity of pre-aspirated stops. *Journal of Linguistics*, 39(03), 575-598.
- Skinner, B. F. ([1957] 2014). *Verbal behavior*. BF Skinner Foundation.

- Smith, L. C., & Baker, W. (2011). L2 dialect acquisition of German vowels: The case of Northern German and Austrian dialects. *Poznań Studies in Contemporary Linguistics PSiCL*, 47, 120-.
- Solon, M., & Kanwit, M. (2014). The Emergence of Future Verbal Morphology in Spanish as a Foreign Language. *Studies in Hispanic & Lusophone Linguistics*, 7(1).
- Sonu, M., Tajima, K., Kato, H., & Sagisaka, Y. (2011). Perceptual training of vowel length contrast of Japanese by L2 listeners: Effects of an isolated word versus a word embedded in sentences. In *Twelfth Annual Conference of the International Speech Communication Association*, 1137-1140.
- Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, 45(3), 412-432.
- Stein, M., Federspiel, A., Koenig, T., Wirth, M., Lehman, C., Wiest, R., Strik, W., Brandeis, D., & Dierks, T. (2009). Reduced frontal activation with increasing 2nd language proficiency. *Neuropsychologia*, 47, 2712–2720.
- Stevens, J. (2011). Vowel duration in second language Spanish vowels: Study abroad versus at-home learners. *Arizona Working Papers in SLA & Teaching*, 18, 77-104.
- Strycharczuk, P., Van 'T Veer, M., Bruil, M., & Linke, K. (2014). Phonetic evidence on phonology–morphosyntax interactions: Sibilant voicing in Quito Spanish. *Journal of Linguistics*, 50(02), 403-452.
- Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language*, 52(3), 322-338.
- Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, 60(4), 487-501.
- Sunderman, G., & Kroll, J. F. (2009). When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics*, 30(1), 79-99.

- Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25(1), 1-24.
- Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning*, 35, 373-404.
- Tarone, E. (2000). Still wrestling with 'context' in interlanguage theory. *Annual Review of Applied Linguistics*, 20, 182-198.
- Tarone, E., & Parish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning: A Journal of Applied Linguistics*, 38(1), 21-44.
- Tarone, E., & Swain, M. (1995). A sociolinguistic perspective on second language use in immersion classrooms. *Modern Language Journal*, 79(2), 166-178.
- Tauroza, S., & Luk, J. (1997). Accent and second language listening comprehension. *RELC Journal*, 28(1), 54-71.
- Terrell, T. D. (1977). Constraints on the aspiration and deletion of final /s/ in Cuban and Puerto Rican Spanish. *The Bilingual Review/La Revista Bilingüe*, 4(1), 35.
- Terrell, T. D. (1978). Sobre la aspiración y elisión de /s/ implosiva y final en el español de Puerto Rico. *Nueva Revista de Filología Hispánica*, 27(1), 24-38.
- Terrell, T. D. (1979). Final /s/ in Cuban Spanish. *Hispania*, 62(4), 599-612.
- Thomas, A. (2004). Phonetic norm versus usage in advanced French as a second language. *International Review of Applied Linguistics in Language Teaching*, 42(4), 365-382.
- Torreira, F. (2006). Coarticulation between aspirated-s and voiceless stops in Spanish: An interdialectal comparison. In N. Sagarra & A. J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, 113-120.

- Torreira, F. (2007). Pre-and post-aspirated stops in Andalusian Spanish. In P. Prieto, J. Mascaró & M. Josep-Solé (Eds.), *Segmental and Prosodic Issues in Romance Phonology*, (67-82). Amsterdam: John Benjamins.
- Torreira, F. (2012). Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association*, 42(01), 49-63.
- Torreira, F., & Ernestus, M. (2012). Weakening of intervocalic /s/ in the Nijmegen Corpus of Casual Spanish. *Phonetica*, 69(3), 124-148.
- Trenchs-Parera, M. (2009). Effects of formal instruction and a stay abroad on the acquisition of native-like oral fluency. *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, 65(3), 365-393.
- Trenchs-Parera, M., & Juan-Garau, M. (2014). A longitudinal study of learners' motivation and beliefs in at home and study abroad contexts. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*, (259-282). Amsterdam: John Benjamins.
- Trubetzkoy, N. S. (1939/1969). *Principles of Phonology*. Berkely, CA: University of California Press.
- Uber, D. R. (1989). Noun-phrase pluralization in the Spanish of Cuban Mariel entrants. *Hispanic Linguistics*, 3, 75-88.
- Uruburu Bidaurrázaga, A. (1990). *Estudios sobre la lengua española en Córdoba*, Diputación Provincial de Córdoba.
- Valdivieso, H., & Magana, J. (1991). Variación fonética de /s/ en el habla espontánea. *Revista de Lingüística Teórica y Aplicada*, 29, 97-113.
- Valls-Ferrer, M. (2011). *The development of oral fluency and rhythm during a study abroad period*. (Unpublished doctoral dissertation). Universitat Pompeu Fabra, Barcelona.

- Valls-Ferrer, M., & Mora, J. C. (2014). L2 fluency development in formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*, (111-136). Amsterdam: John Benjamins.
- VandeBerg, M., Connor-Linton, J., & Paige, R. M. (2009). The Georgetown Consortium Project: Interventions for student learning abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 18, 1-75.
- Van Hell, J. G., & Tanner, D. (2012). Second Language Proficiency and Cross-Language Lexical Activation. *Language Learning*, 62(s2), 148-171.
- Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39(3), 458-483.
- VanPatten, B. (2004). *Input processing in second language acquisition. Processing instruction: Theory, research, and commentary*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Veguez, R. (April, 1984). The Oral Proficiency Interview and the junior year abroad: Some unexpected results. *Paper presented at the Annual Northeast Conference on the Teaching of Foreign Languages*, New York, NY.
- Villena-Ponsoda, J. A. (2008). Sociolinguistic patterns of Andalusian Spanish. *International Journal of the Sociology of Language*, 193/194, 139-160.
- Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, 113(2), 1033-1043.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49-63.

- Whatley, M. (2013). The acquisition of past tense variation by L2 Learners of Spanish in an abroad context. In J. Cabrelli Amaro, G. Lord, A. de Prada Pérez & J. E. Aaron (Eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, 190-205. Somerville, MA: Cascadilla Proceedings Project.
- Widdison, K. A. (1995a). Two models of Spanish s-aspiration. *Language Sciences*, 17(4), 329-343.
- Widdison, K. A. (1995b). An acoustic and perceptual study of the Spanish sound change s > h. *Rivista di Linguistica*, 7, 175-190.
- Widdison, K. A. (1997). Phonetic explanations for sibilant patterns in Spanish. *Lingua*, 102(4), 253-264.
- Willis, E., Geeslin, K., & Henriksen, N. (October, 2009). The acquisition of /θ/ by study abroad learners in León, Spain. *Paper presented at the 13th Hispanic Linguistics Symposium*, San Juan, Puerto Rico.
- Wilkinson, S. (1996). Foreign language conversation and the study abroad transition: A case study. *Dissertation Abstracts International*, 57, 140-A.
- Wilkinson, S. (1998). On the nature of immersion during study abroad: Some participant perspectives. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 1, 121-138.

Appendix A: Grammar test

Instructions: Please read the story below about a Hispanic female college student and select the answers that best complete each sentence.

Creo que es muy interesante _____ (**hablando;hablo;hablar**) de los hábitos alimenticios de la gente. Yo, por mi parte, soy vegetariana. Cuando voy a eventos sociales, como por ejemplo fiestas, bodas o bailes, espero que _____ (**hay;sea;haya**) comida vegetariana allí. Algunas personas dicen que _____ (**los;le;les**) representa un inconveniente proveer _____ (**la;lo;le**) , pero yo creo que no _____ (**tiene;tengo;tenga**) que ser así. De hecho, la comida vegetariana es muy fácil _____ (**a;en;de**) preparar. Y cuando no se ofrece, puede ser _____ (**una;el;un**) gran problema. Yo recuerdo una vez que _____ (**fui;voy;iba**) a una fiesta de cumpleaños y _____ (**resultaba;resulté;resultó**) ser todo un desastre. La fiesta era en la casa de un amigo, y él había invitado a mucha gente. Me sorprendió porque para ser un estudiante de postgrado con poco dinero, tenía una gran variedad de comida para los invitados. Yo creo que si me _____ (**habría;había;hubiera**) tocado a mí dar la fiesta, no _____ (**habría;había;hubiera**) dado ni la mitad de lo que _____ (**había;hubiera;era**) allí. Pero pronto me _____ (**doy;daba;di**) cuenta que él no había preparado nada vegetariano. Yo no pongo problemas por ese tipo de cosas, pero una amiga _____ (**mía;de mí;mi**) sí _____ (**le;se;lo**) hace. _____ (**Empezado;Empezaba;Empezó**) a quejarse en frente de todo el mundo, mientras el anfitrión sólo _____ (**miraría;miraba;miró**) la escena con _____ (**una;la;su**) boca abierta. Yo le dije a mi amiga que _____ (**dejaba;deje;dejara**) de causar tanto escándalo, pero no me puso atención. Por fin, el anfitrión dijo: “La próxima vez que tenga una fiesta, _____ (**preparara;prepararía;prepararé**) algo vegetariano.” Yo le dije después a mi amiga: “Mejor tarde que nunca, ¿no?”

I think it is very interesting _____ (**speaking;I speak;to speak**) of the eating habits of people. I, for one, am vegetarian. When I go to social events, like for example parties, weddings or dances, I hope that _____ (**there is-indicative;there be;there is-subjunctive**) vegetarian food there. Some people say that _____ (**them;to him/her;to them**) it is problem to provide _____ (**it-FEM;it-MASC;to him/her**), but I believe that it doesn't _____ (**have-3sg;have-1sg;have-3sg subjunctive**) to be that way. In fact, vegetarian food is very easy _____ (**to;on/in;of/from**) to prepare. And when it is not offered, it can be _____ (**a-FEM;the-MASC;a-MASC**) big problem. I remember a time that _____ (**I went;I go;I was going**) to a birthday party and _____ (**it was ending up;I ended up;it ended up**) being a total disaster. The party was at a friend's house, and he had invited a lot of people. It surprised me because, for being a graduate student with little money, he had a large variety of food for the guests. I believe that if _____ (**it would have;had;had-subjunctive**) been my turn to throw the party, I _____ (**would not;had not;had not-subjunctive**) given even half of what _____ (**there was;there was-subjunctive;was**) there. But soon I _____ (**realize;was realizing;realized**) that he hadn't prepared anything vegetarian. I don't care about those things but a friend _____ (**3 forms of possessive mía;de mí;mi**) yes, does _____ (**IndObj;ReflPron;DirObj**). _____ (**Begun;She was beginning;She began**) to complain in front of everyone, while the host just _____ (**would watch;was watching;watched**) the scene with _____ (**IndefArt-Fem;DefArt-Fem;PossAdj**) mouth open. I told my friend that _____ (**she was stopping;she stop-present;she stop-past**) causing such a scandal, but she didn't pay any attention to me. Finally, the host said: “The next time I have a party, _____ (**prepared-ImpSubjunctive;I would prepare; I will prepare**) something vegetarian.” I told my friend afterward: “Better late than never, right?”

Appendix B: Example of Identification Task Screenshot

Figure 23. Praat screenshot of the identification response options for the stimulus <gasco> ([gak^ho])



Appendix C: Acoustic Measurements of the Identification Task Target Stimuli (in milliseconds)

Coda Condition: No Coda					
TARGET WORD	DURATION OF FIRST VOWEL	ASPIRATION DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
bape	122.50	0.00	16.14	118.06	122.50
bipa	115.87	0.00	9.16	116.49	115.87
cheto	111.02	0.00	11.45	142.98	111.02
daca	139.04	0.00	37.48	113.01	139.04
depa	151.69	0.00	11.52	115.54	151.69
dipe	130.88	0.00	15.47	115.65	130.88
feca	104.68	0.00	25.73	108.34	104.68
feco	129.53	0.00	38.82	116.21	129.53
fepe	179.16	0.00	18.94	128.65	179.16
fite	103.89	0.00	18.49	123.59	103.89
fote	139.29	0.00	22.57	129.04	139.29
gaco	127.25	0.00	18.05	94.02	127.25
gapo	129.55	0.00	22.89	123.09	129.55
goco	177.14	0.00	25.14	107.15	177.14
gote	123.53	0.00	27.74	110.37	123.53
jata	163.81	0.00	14.64	126.61	163.81
lepo	123.37	0.00	28.62	124.76	123.37
lete	232.32	0.00	10.10	136.48	232.32
lico	110.21	0.00	46.41	95.17	110.21
meque	124.43	0.00	27.43	118.54	124.43
mipo	97.58	0.00	9.26	130.83	97.58
nique	158.85	0.00	38.53	101.78	158.85
nita	124.27	0.00	16.97	112.30	124.27
plito	106.19	0.00	12.97	130.84	106.19
queto	127.25	0.00	30.57	133.42	127.25
quique	97.74	0.00	20.73	117.52	97.74
soca	116.73	0.00	24.16	104.75	116.73
tepo	85.68	0.00	10.16	136.51	85.68

Coda Condition: Post-aspirated

TARGET WORD	DURATION OF FIRST VOWEL	ASPIRATION DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
bap^he	146.01	34.39	39.62	188.48	180.40
bip^ha	86.40	7.76	94.71	126.39	94.16
chet^ho	100.72	6.15	46.56	146.62	106.86
dac^ha	110.06	11.72	55.33	123.09	121.78
dep^ha	111.88	37.86	38.06	130.42	149.73
dip^he	91.35	9.60	49.58	133.99	100.95
fec^ha	85.02	22.73	58.01	119.32	107.75
fek^ho	87.87	19.56	64.40	138.67	107.43
fep^he	108.20	28.01	44.24	165.96	136.20
fit^he	90.88	7.03	42.37	146.62	97.91
fot^he	114.94	20.02	31.08	126.92	134.96
gac^ho	131.41	23.17	61.05	113.31	154.58
gap^ho	131.09	47.60	40.03	158.73	178.69
goc^ho	135.93	43.55	75.80	94.04	179.48
got^he	100.69	3.90	32.52	140.80	104.59
jat^ha	83.40	23.36	53.99	108.66	106.77
lep^ho	108.16	33.19	65.78	170.10	141.35
let^he	127.39	17.04	47.82	168.32	144.42
lic^ho	97.85	26.45	92.09	113.71	124.30
mequ^he	100.05	24.62	58.59	129.47	124.67
mip^ho	122.25	25.56	56.45	144.83	147.81
niqu^he	98.67	27.51	56.63	134.71	126.17
nit^ha	113.76	34.07	60.42	156.86	147.83
plit^ho	88.07	44.10	46.84	110.38	132.16
quet^ho	83.87	42.17	51.25	172.96	126.04
quiqu^he	84.28	19.61	65.26	123.77	103.89
soc^ha	103.54	25.40	60.39	111.13	128.94
tep^ho	130.95	16.93	58.39	157.19	147.88

Coda Condition: Sibilant

TARGET WORD	DURATION OF FIRST VOWEL	SIBILANT DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + SIBILANT DURATION
baspe	125.49	82.37	6.46	104.07	207.86
bispa	93.04	84.49	17.33	118.66	177.53
chesto	93.32	66.96	20.53	98.13	160.28
dasca	129.04	90.34	28.55	99.85	219.38
despa	112.58	73.46	5.96	109.34	186.05
dispe	95.18	95.42	11.45	98.54	190.59
fesca	99.79	98.99	21.27	62.83	198.78
fesco	108.97	93.05	25.56	91.28	202.01
fespe	152.36	107.17	14.11	112.37	259.54
fiste	128.81	133.78	18.16	96.99	262.59
foste	139.69	89.24	21.58	110.35	228.93
gasco	137.78	85.61	25.77	86.64	223.39
gaspo	163.18	100.46	22.89	92.19	263.64
gosco	122.49	93.69	28.45	80.38	216.18
goste	162.48	79.46	19.20	88.52	241.94
jasta	107.62	87.89	13.97	92.41	195.50
lespo	123.68	84.41	20.52	117.13	208.09
leste	147.36	82.31	17.50	113.27	229.67
lisco	114.92	89.29	22.50	101.11	204.21
mesque	107.16	100.00	19.36	74.76	207.16
mispo	107.34	94.47	6.17	110.11	201.81
nisque	114.77	86.46	16.16	81.33	201.23
nista	150.35	107.93	20.61	94.16	258.27
plisto	94.36	89.37	27.53	109.82	183.73
questo	131.90	89.87	27.77	89.76	221.77
quisque	122.86	109.86	23.57	89.21	232.73
sosca	106.81	93.60	18.33	113.28	200.42
tespo	118.95	86.78	15.46	88.30	205.73

Appendix D: Waveforms and Spectrograms of Stimuli Representing the Coda Conditions from the Identification Task (Male talkers)

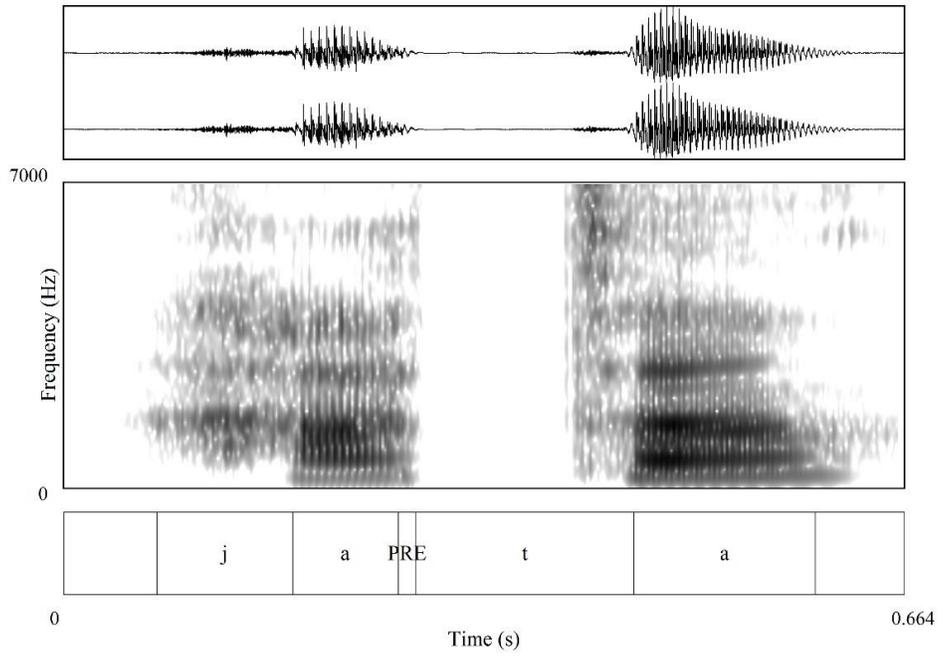


Figure 24: Post-aspiration, *jasta* [jat^ha]

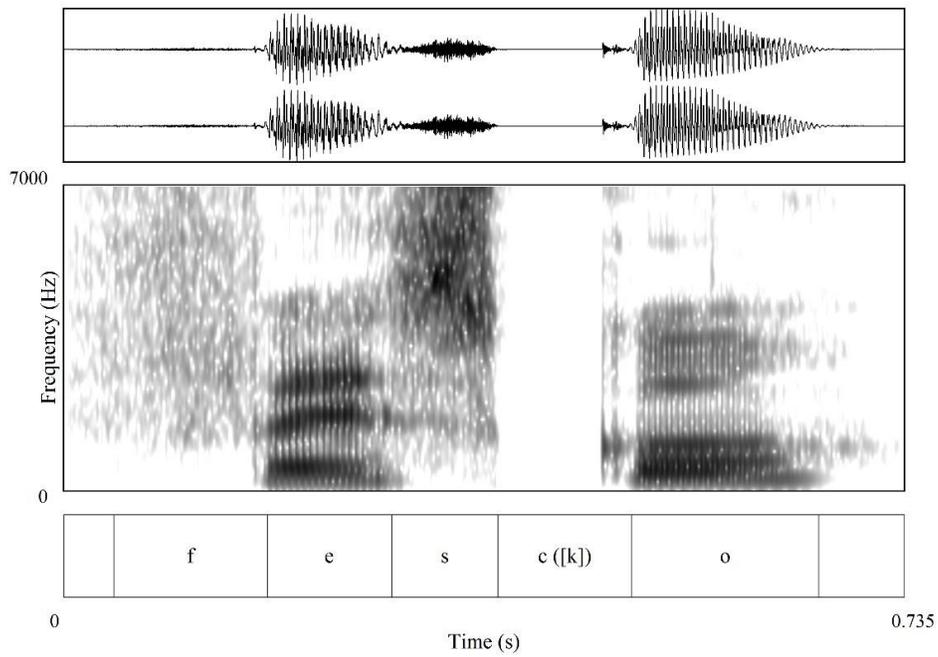


Figure 25: Sibilant [s], *fesco* [fesko]

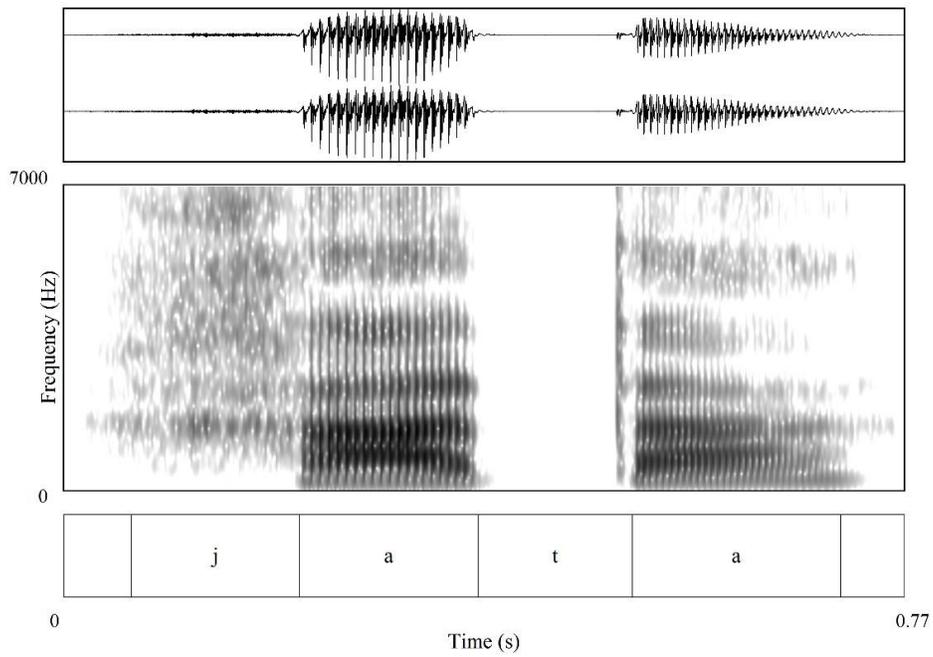


Figure 26: No coda, jata [xata]

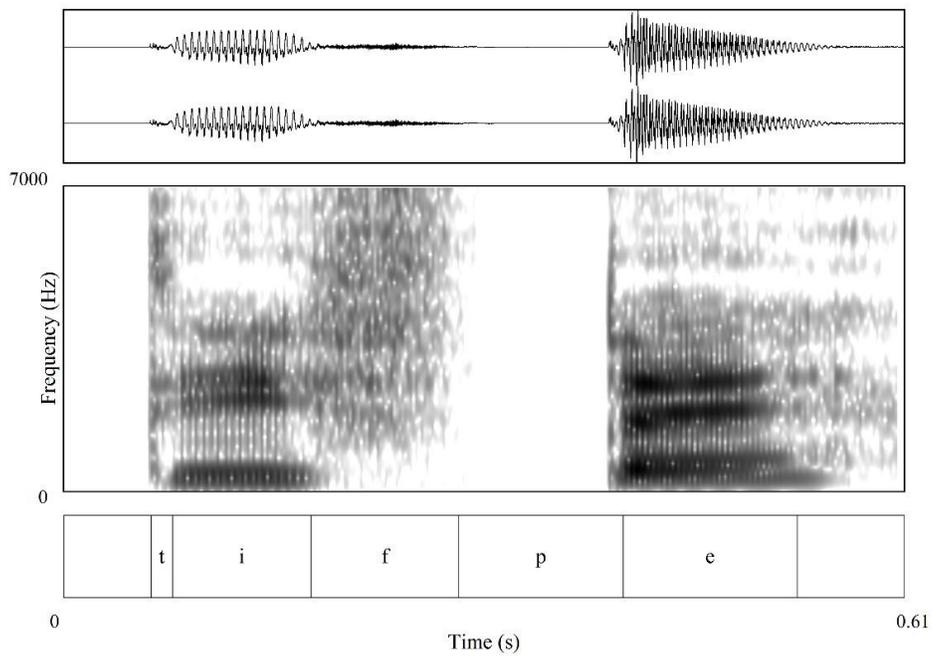


Figure 27: [f], tife [tifpe]

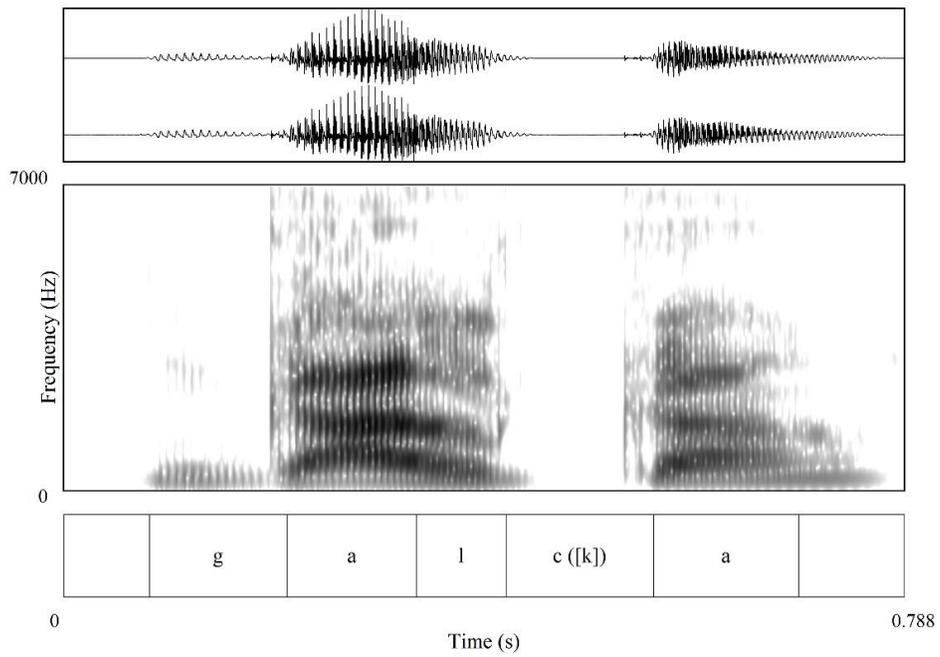


Figure 28: [l], galca [galka]

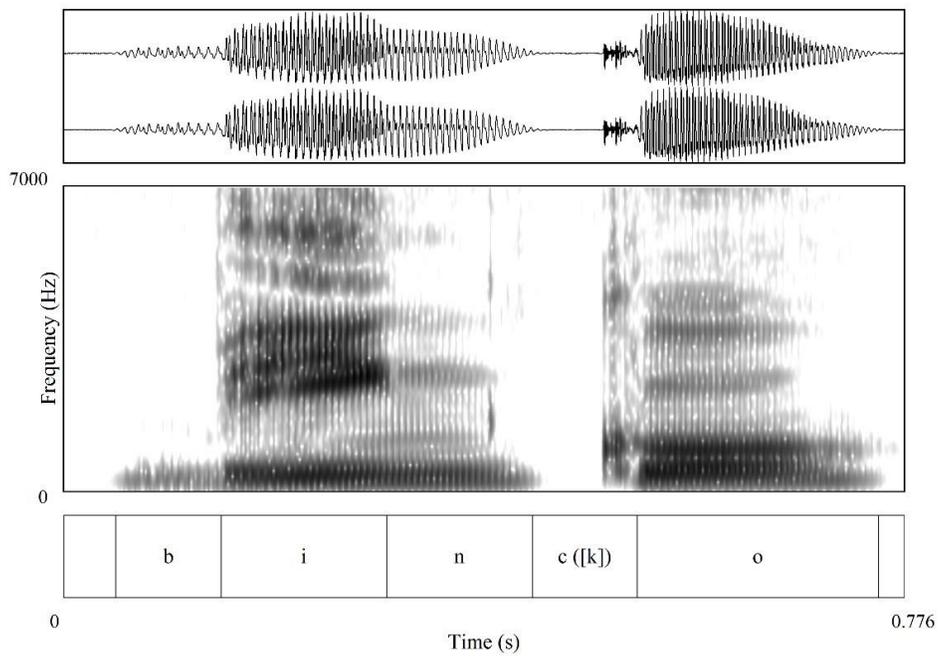


Figure 29: [n], binco [biŋko]

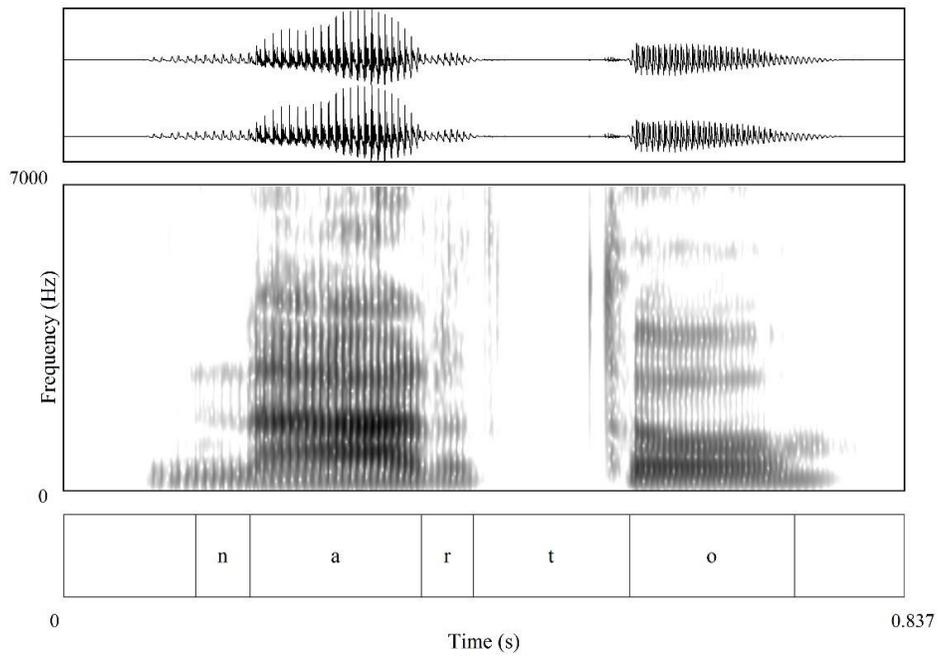


Figure 30: [r], narto [narto]

Appendix E: Acoustic Measurements of the Lexical Decision Task Target Stimuli

Coda Condition: Sibilant

TARGET WORD	REAL/NON-WORD STATUS	DURATION OF FIRST VOWEL	SIBILANT DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
amistad	Real word	72.42	121.62	12.69	73.76	194.04
amostad	Non-word	78.35	55.80	11.11	68.72	134.15
artista	Real word	120.52	138.95	22.57	66.69	259.47
ascucha	Non-word	93.94	84.27	37.29	68.48	178.21
asiste	Real word	90.29	76.71	13.83	94.93	167.00
aspecto	Real word	82.28	97.80	9.56	122.40	180.08
aspicto	Non-word	103.68	97.67	11.65	89.62	201.35
aspira	Real word	73.25	94.04	8.36	86.54	167.29
astilo	Non-word	113.98	82.84	24.20	76.99	196.83
bisque	Non-word	86.71	103.72	17.43	76.50	190.43
bosque	Real word	128.85	88.81	18.42	89.99	217.66
busca	Real word	111.35	86.96	19.53	92.19	198.31
costambre	Non-word	92.81	73.12	8.26	79.49	165.94
daspués	Non-word	62.05	68.03	52.29	91.36	130.08
descanso	Real word	82.08	71.41	20.64	73.33	153.49
después	Real word	92.45	70.04	12.47	129.65	162.50
destimo	Non-word	84.29	74.30	18.23	75.97	158.59
destino	Real word	130.17	114.38	22.21	71.16	244.56
discanso	Non-word	79.87	91.94	25.51	75.42	171.81
disco	Real word	112.25	90.57	22.07	82.12	202.82
discurso	Real word	88.97	81.22	38.32	80.20	170.19
discute	Real word	85.04	93.36	19.56	72.90	178.40
dusto	Non-word	124.66	64.83	16.34	97.28	189.50
escapa	Real word	85.64	101.32	18.27	99.33	186.96
escopa	Non-word	100.14	78.38	27.89	88.18	178.52
escribir	Real word	72.24	68.23	25.75	56.25	140.48
escripir	Non-word	73.40	74.05	14.87	80.99	147.45
escucha	Real word	85.74	75.64	25.77	80.42	161.38
espacio	Real word	70.88	86.43	7.36	117.31	157.31
espajo	Non-word	84.58	119.77	7.06	133.13	204.35
espejo	Real word	127.23	110.46	8.82	112.17	237.69
espera	Real word	96.79	110.63	9.18	92.39	207.42
espisa	Non-word	96.90	100.45	11.51	93.21	197.35
esposa	Real word	78.78	78.18	16.81	89.79	156.96
esta	Real word	122.41	68.68	13.68	99.31	191.09
estado	Real word	109.64	79.33	16.28	61.43	188.97

TARGET WORD	REAL/NON-WORD STATUS	DURATION OF FIRST VOWEL	SIBILANT DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
éste	Real word	86.25	84.05	19.62	120.77	170.31
estedo	Non-word	84.47	67.57	11.99	88.83	152.04
estilo	Real word	94.21	72.45	20.34	76.36	166.66
esto	Real word	108.96	116.11	14.28	120.70	225.07
fiesta	Real word	160.27	59.53	14.24	98.49	219.80
fispano	Non-word	63.96	83.38	7.98	88.39	147.35
fresco	Real word	115.88	84.99	39.88	94.89	200.87
frisco	Non-word	91.64	77.11	32.80	84.35	168.75
gosto	Non-word	140.47	76.15	11.79	72.22	216.62
gusto	Real word	110.68	80.26	17.54	68.58	190.94
hispano	Real word	88.80	88.97	10.22	111.76	177.77
homesto	Non-word	119.57	74.82	17.78	87.58	194.39
honesto	Real word	134.69	65.44	20.93	99.01	200.13
ispacio	Non-word	74.25	109.48	7.98	91.12	183.73
ispera	Non-word	74.14	90.96	15.03	88.18	165.09
ista	Non-word	132.77	126.08	18.96	75.78	258.84
iste	Non-word	113.18	133.80	27.81	97.27	246.98
isto	Non-word	104.07	139.93	18.48	85.73	244.00
jiesta	Non-word	155.20	106.19	19.15	87.06	261.39
justo	Real word	108.32	85.34	12.18	98.36	193.66
lesponde	Non-word	64.25	73.77	7.42	95.64	138.02
nespeto	Non-word	80.49	63.64	10.15	103.08	144.13
oaste	Non-word	103.44	90.10	16.14	90.19	193.54
oeste	Real word	111.04	104.43	12.66	85.22	215.47
opuesto	Real word	141.18	94.81	26.84	68.38	235.99
ortista	Non-word	88.62	65.30	25.16	81.76	153.92
oscuro	Real word	114.08	90.12	44.48	96.48	204.21
osiste	Non-word	87.22	77.01	13.13	89.04	164.23
ospira	Non-word	88.52	73.78	13.04	117.56	162.29
osquero	Non-word	80.94	103.48	13.85	80.21	184.42
osted	Non-word	63.74	67.74	16.06	64.63	131.49
piscurso	Non-word	89.75	109.64	35.37	96.23	199.38
piscute	Non-word	77.37	88.19	26.79	98.51	165.56
presta	Real word	96.12	82.71	20.38	103.10	178.82
prista	Non-word	96.02	140.70	29.65	45.24	236.73
puasto	Non-word	180.85	96.51	25.33	109.02	277.36
puesto	Real word	123.13	86.97	15.40	76.55	210.10
pusca	Non-word	99.32	77.01	32.31	85.40	176.33

TARGET WORD	REAL/NON-WORD STATUS	DURATION OF FIRST VOWEL	SIBILANT DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
ravista	Non-word	122.71	81.20	6.81	108.87	203.91
respeto	Real word	125.85	102.81	12.52	92.14	228.67
responde	Real word	66.34	67.21	11.91	78.67	133.55
revista	Real word	117.26	77.14	16.69	84.39	194.40
sestema	Non-word	83.34	103.81	12.94	68.92	187.15
sistema	Real word	83.63	103.32	14.79	74.02	186.95
sumuesto	Non-word	156.43	88.50	19.46	90.08	244.93
supuesto	Real word	115.15	61.60	13.58	107.95	176.75
tisco	Non-word	115.53	75.88	33.66	73.68	191.41
trista	Non-word	98.84	72.38	21.22	81.57	171.22
triste	Real word	114.94	85.52	13.98	88.90	200.45
upuesto	Non-word	142.35	104.73	26.36	83.07	247.08
usted	Real word	71.70	96.64	11.41	67.55	168.34
usted	Real word	76.73	113.58	20.57	92.22	190.31

Coda Condition: Post-aspirated

TARGET WORD	REAL/NON-WORD STATUS	DURATION OF FIRST VOWEL	ASPIRATION DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
amistad	Real word	52.08	0.00	52.93	116.95	52.08
amostad	Non-word	56.19	19.69	44.10	118.06	75.88
artista	Real word	92.07	16.89	61.99	91.08	108.96
ascucha	Non-word	101.87	47.19	74.20	101.45	149.05
asiste	Real word	110.52	0.00	49.67	171.06	110.52
aspecto	Real word	45.84	20.48	27.45	159.16	66.32
aspicto	Non-word	51.99	39.32	25.52	178.92	91.31
aspira	Real word	62.43	0.00	34.34	128.11	62.43
astilo	Non-word	22.66	56.03	41.47	163.27	78.69
bisque	Non-word	87.13	27.96	53.95	132.55	115.09
bosque	Real word	102.74	0.00	62.99	125.53	102.74
busca	Real word	85.33	17.28	53.20	116.26	102.61
costambre	Non-word	64.74	5.31	35.56	114.64	70.06
costumbre	Real word	52.10	15.73	29.55	94.25	67.83
después	Non-word	82.44	0.00	53.56	144.93	82.44
descanso	Real word	76.58	15.32	51.57	88.04	91.90

TARGET WORD	REAL/NON-WORD STATUS	DURATION OF FIRST VOWEL	ASPIRATION DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
después	Real word	74.72	9.48	53.38	89.73	84.20
destimo	Non-word	64.90	26.28	29.48	125.87	91.18
destino	Real word	80.76	8.37	52.08	120.86	89.13
discanso	Non-word	79.62	0.00	67.35	112.18	79.62
disco	Real word	82.67	31.71	59.69	107.12	114.37
discurso	Real word	50.48	30.58	86.46	75.69	81.06
discute	Real word	52.79	25.36	90.25	78.02	78.15
dusto	Non-word	106.20	0.00	40.42	159.71	106.20
escapa	Real word	61.48	18.07	46.20	94.14	79.54
escopa	Non-word	65.05	24.84	57.17	107.36	89.89
escribir	Real word	48.03	20.71	83.64	83.64	68.74
escripir	Non-word	85.53	0.00	51.00	102.79	85.53
escucha	Real word	89.95	44.65	71.47	48.43	134.60
espacio	Real word	90.33	0.00	30.06	128.35	90.33
espajo	Non-word	45.90	32.10	20.56	146.66	77.99
espejo	Real word	56.61	21.27	25.65	144.11	77.88
espera	Real word	85.43	0.00	40.58	134.64	85.43
espisa	Non-word	78.99	21.08	41.72	170.76	100.07
esposa	Real word	73.41	42.19	20.02	130.71	115.61
esta	Real word	91.51	16.58	41.81	156.92	108.10
estado	Real word	87.99	0.00	25.86	122.34	87.99
éste	Real word	98.52	0.00	50.12	158.71	98.52
estedo	Non-word	71.75	18.05	28.18	141.12	89.80
estilo	Real word	97.81	0.00	42.86	100.84	97.81
esto	Real word	89.94	31.41	50.71	160.43	121.34
fiesta	Real word	112.07	24.26	43.66	132.43	136.34
fispano	Non-word	75.89	0.00	30.30	123.04	75.89
fresco	Real word	70.06	25.82	73.36	99.68	95.88
frisco	Non-word	77.82	25.80	78.35	90.03	103.62
gosto	Non-word	112.67	34.51	31.84	138.38	147.18
gusto	Real word	124.32	17.82	53.92	92.15	142.14
hispano	Real word	68.18	21.75	45.91	117.09	89.93
homesto	Non-word	104.65	0.00	40.66	160.54	104.65
honesto	Real word	107.84	17.14	38.07	137.28	124.98
ispacio	Non-word	81.53	0.00	38.99	83.35	81.53
ispera	Non-word	59.40	28.53	47.92	132.87	87.93
ista	Non-word	87.28	0.00	50.15	165.38	87.28
iste	Non-word	111.00	0.00	46.43	134.16	111.00

TARGET WORD	REAL/NON-WORD STATUS	DURATION OF FIRST VOWEL	ASPIRATION DURATION (PRECEDING THE STOP)	VOT DURATION	STOP CLOSURE DURATION	FIRST VOWEL + ASPIRATION DURATION
isto	Non-word	98.71	25.05	57.83	171.05	123.76
jiesta	Non-word	146.42	33.08	57.95	97.68	179.50
justo	Real word	99.91	0.00	49.58	155.21	99.91
lesponde	Non-word	65.30	27.53	32.33	104.69	92.83
nespeto	Non-word	86.09	33.45	29.04	138.31	119.54
oaste	Non-word	111.55	0.00	36.22	155.08	111.55
oeste	Real word	91.42	16.56	59.01	153.71	107.98
opuesto	Real word	144.12	0.00	58.08	135.66	144.12
ortista	Non-word	94.82	23.70	44.50	106.89	118.52
oscuro	Real word	83.28	31.59	67.69	89.50	114.87
osiste	Non-word	81.88	21.58	39.50	131.52	103.46
ospira	Non-word	92.52	32.98	36.44	124.24	125.50
osquero	Non-word	78.92	0.00	58.82	110.25	78.92
osted	Non-word	71.63	16.06	45.29	133.33	87.68
piscurso	Non-word	58.61	20.52	48.15	119.82	79.13
piscute	Non-word	70.55	15.58	70.60	106.55	86.13
presta	Real word	94.39	29.78	39.74	136.45	124.18
prista	Non-word	94.47	22.26	41.16	112.39	116.73
puasto	Non-word	126.07	28.65	45.81	144.53	154.72
puesto	Real word	115.55	13.04	39.64	133.41	128.59
pusca	Non-word	86.00	53.74	51.60	100.85	139.74
ravista	Non-word	87.92	31.86	36.23	157.97	119.78
respeto	Real word	86.60	15.08	41.17	111.75	101.68
responde	Real word	49.27	28.40	46.98	110.72	77.66
revista	Real word	115.72	0.00	46.48	137.89	115.72
sestema	Non-word	69.19	11.96	42.38	105.51	81.16
sistema	Real word	48.82	9.99	33.88	111.45	58.81
sumuesto	Non-word	133.10	0.00	41.30	152.23	133.10
supuesto	Real word	91.60	15.41	32.44	149.30	107.01
tisco	Non-word	102.21	40.23	58.15	86.70	142.44
trista	Non-word	99.18	0.00	44.60	145.87	99.18
triste	Real word	100.48	18.17	25.79	116.88	118.65
upuesto	Non-word	134.25	21.35	47.88	168.85	155.61
usted	Real word	65.19	21.40	52.63	110.19	86.58

Appendix F: Waveforms and spectrograms comparing the target lexical coda conditions

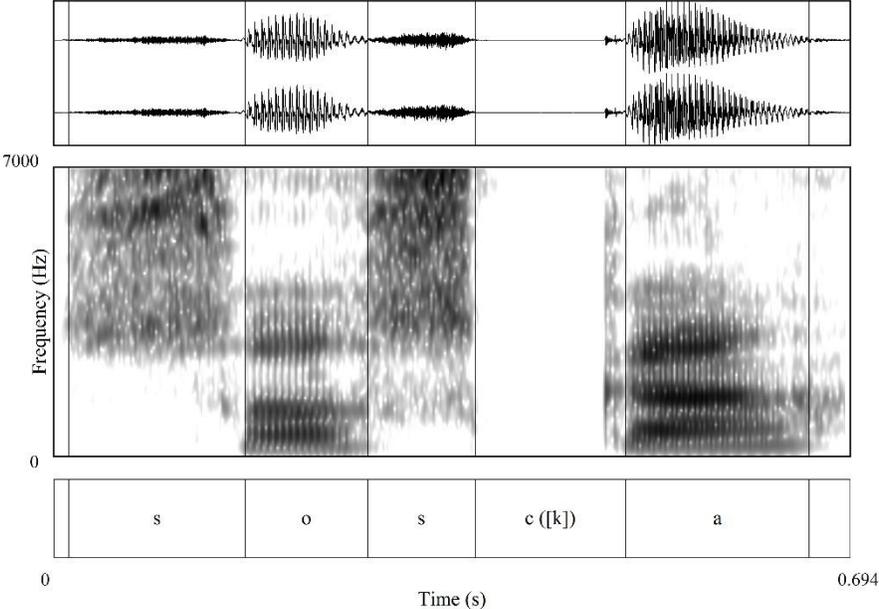


Figure 31: Example of a CodaS condition, *sosca* [soska]

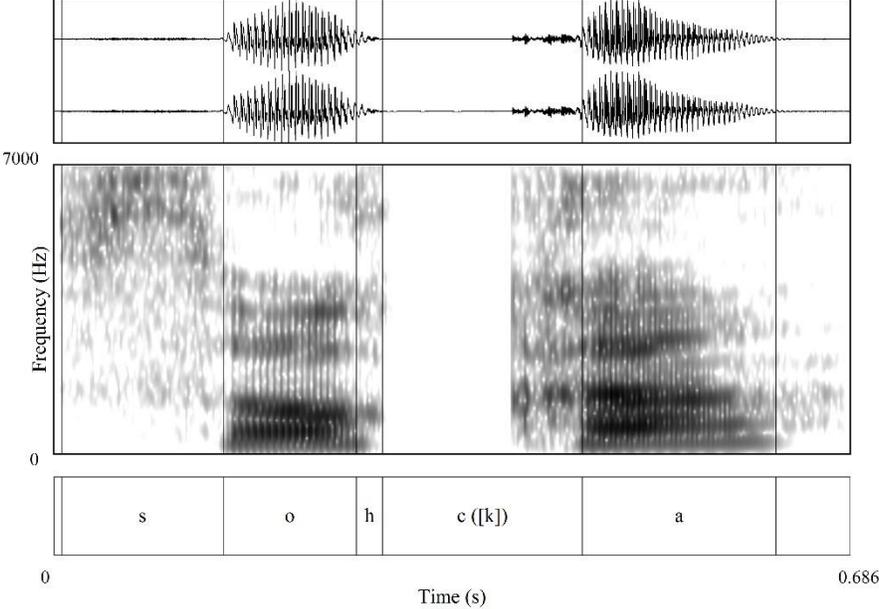


Figure 32: Example of the Asp condition, *sosca* [sok^ha]

Appendix G: Language Background Questionnaire (i.e., Time 1 questionnaire – study abroad groups)

Please enter your participant code accurately.

You are

- Male
- Female

What is your email address?

What is the name of your home university in the U.S.?

Through which organization is your study abroad program in Seville offered?

- [Study Abroad 1] (to maintain the anonymity of the programs, the names were removed)
- [Study Abroad 2]

In which program will you study this semester? (only for Study Abroad 2 students)

- International Business and Culture
- Business and Society
- Liberal Arts
- Advanced Liberal Arts
- Language and Society
- Communication, New Media, and Journalism

Please select the native language(s) that you and your parents speak (select all that apply). For this question, a native language is a language that a person is exposed to from birth and grows up speaking.

	English	Spanish	Portuguese	French	Italian	Other/Not applicable
My native language(s)	<input type="checkbox"/>					
Mother's native language(s)	<input type="checkbox"/>					
Father's native language(s)	<input type="checkbox"/>					

You selected "Other" at least once on the previous question. Please name the other language(s) and to which person(s) it applies in the space below:

Which of these languages is/are spoken consistently among your immediate family members (select all that apply)

- English
- Spanish
- Portuguese
- French
- Italian
- Other (Please list languages here) _____

What level of Spanish courses will you be taking this semester?

- Beginner
- Intermediate
- Advanced
- I don't know yet
- I will only be taking courses taught in English

Which of the following best describes your major/minor at your home university?

- Spanish major
- Spanish minor
- Neither

What is/are your major(s) and minor(s)?

How many years have you studied Spanish in the following levels of education?

	N/A	Less than 1 year	1 to 2 years	3 or more years
Elementary School	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Middle/Junior High School	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
High School	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Select the type of Spanish courses which you have completed at each level of university study (mark all that apply). Note: A Spanish linguistics course is any course that has the purpose of discussing the scientific study of Spanish as a language and its parts (speech sounds, word structure, sentence structure, etc.). It is not the same as a grammar course.

	Grammar	Conversation	Composition (focus on writing)	Literature	Linguistics	Culture	Other
First year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Second year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Third year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fourth year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Graduate level course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Have you ever attended a school where Spanish was the primary language of instruction? (e.g. Spanish immersion school) If yes, please explain when and for how long.

- Yes _____
- No

Previous to this study abroad, had you ever visited/lived in a Spanish-speaking country for 3 weeks or more at one time?

- Yes
- No

Please provide the name of each Spanish-speaking country and city/cities you have visited/lived in for 3 weeks or more along with: (1) the date(s) you were there (e.g. 2005-2006 if full year or July 2008-September 2008) (2) the duration of the stay in months and weeks (e.g. 3 months and 2 weeks) (3) your primary purpose(s) for living there (e.g. language study, vacation, service, Peace Corps, military, mission etc.) If you have been to the same place multiple times for 3 weeks or longer, use separate lines for each time.

	Date(s)	Duration	Purpose(s)
Country and city/cities #1			
Country and city/cities #2			
Country and city/cities #3			
Country and city/cities #4			
Country and city/cities #5			
Country and city/cities #6			
Country and city/cities #7			
Country and city/cities #8			

In the last year before arriving in Seville, on average, how often did you participate outside of class in the following activities in Spanish?

	A few times or less per year	Once a month	A few times per month	Once per week	A few times or more per week
Speaking Spanish with native speakers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking Spanish with non-native speakers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Watching TV, videos, or movies in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading articles or books in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Listening to music in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Communicating online in Spanish (Facebook, Twitter, email etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texting in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have had any relatively frequent contact with native Spanish speakers within the past 2 years, please complete the following information about them and about your contact with them: Under the far left-hand column ("Native speaker #") put the sex (M/F), age (estimate if unknown) of each person (e.g., M, 24)

	Country of origin	City of origin (if known)	Approximate frequency of contact with this person in Spanish, either listening to them or speaking to them (daily, weekly, monthly, annually)	Relationship to you (family, friend, co-worker, teacher, acquaintance)
Native speaker #1				
Native speaker #2				
Native speaker #3				
Native speaker #4				
Native speaker #5				
Native speaker #6				

In your opinion, how do your skills in Spanish compare to the other second-language Spanish learners in your Spanish classes?

	Below Average	Average	Above Average	Well Above Average
Understanding spoken speech	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking (speed of speech, hesitations, pauses, & overall errors)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pronunciation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grammar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Have you studied/learned any other languages other than Spanish?

- Yes
- No

Please list the language(s) on the left and select your approximate overall proficiency in each language on the right.

Language 1	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced
Language 2	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced
Language 3	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced

Please rank your reasons for studying abroad based on importance (only one reason can be chosen or each number 1-9) 1 = most important, 9 = least important

- _____ Experience a foreign culture
- _____ Expand my circle of friends
- _____ Explore my heritage
- _____ Prepare for a future career in which I would use Spanish
- _____ Fulfill academic requirements
- _____ Improve my Spanish
- _____ See the world/travel around
- _____ Build my résumé to be more competitive in the job market
- _____ I didn't have a choice

Please rank the following language learning priorities in the order of how important they are to you this semester as you try to improve your Spanish: 1 = most important, 7 = least important

- _____ Learn how to speak with the local accent
- _____ Improve my overall pronunciation
- _____ Improve my overall grammar
- _____ Learn how to read better in Spanish
- _____ To better understand native speakers when they speak
- _____ Learn how to write better in Spanish
- _____ To better understand the cultural aspects (e.g., appropriateness, politeness etc.) of the Spanish language

Please rank the following activities in order of their priority to you during this semester abroad: 1 = most important, 8 = least important

- _____ Studying/School work
- _____ Making friends with native Spanish speakers and spending time with them
- _____ Local sightseeing around Seville
- _____ Visiting other places outside of Seville
- _____ Making friends with English-speaking students from my program and spending time with them
- _____ Trying new cultural experiences (e.g., new foods, new activities)
- _____ Participating in some sort of service learning in the local community
- _____ Partying and/or Night scene

Please use the slider to rate how strongly you agree or disagree with the following statements. Moving the slider farther to the right indicates a stronger agreement and farther to the left indicates a stronger disagreement. The middle (50) is neutral. Please answer honestly. Your answers will not be shared with anyone else with your name attached to them.

- _____ I love the Spanish language
- _____ I mainly study Spanish to build my résumé
- _____ Learning Spanish is boring to me
- _____ Language learning is a fun and interesting process
- _____ I am studying abroad primarily to fulfill a language requirement at my university
- _____ I am studying abroad primarily to improve my Spanish language skills and experience the language in its cultural context
- _____ For me, studying abroad in Seville is more about having fun than about learning Spanish
- _____ I am excited to take Spanish classes while in Seville
- _____ I want to learn how to speak like the people from Seville while I am here
- _____ I am not interested in learning to speak a new dialect of Spanish
- _____ I care a lot about how "native-like" my Spanish sounds
- _____ I care more about getting my point across than how my Spanish sounds

The reason(s) I chose to study abroad in Seville as opposed to other Spanish-speaking places is because...(mark all that apply)

- It was more economical
- I prefer the courses offered
- I was encouraged by someone to study in Seville
- I find Seville's culture more interesting than others
- Seville is more exotic/interesting than other Spanish-speaking places
- I assumed the local people would be friendly
- I think the dialect of Spanish spoken in Seville will be interesting
- I want to learn to speak Spanish like people from Seville
- Other _____

Do you have any history of trouble with hearing?

- Yes
- No

Appendix H: Language Background Questionnaire (Time 1 - At-home group)

Please enter your participant code accurately (in all caps)

You are

- Male
- Female

What is your email address?

Please select the native language(s) that you and your parents speak (select all that apply). For this question, a native language is a language that a person is exposed to from birth and grows up speaking.

	English	Spanish	Portuguese	French	Italian	Other/Not applicable
My native language(s)	<input type="checkbox"/>					
Mother's native language(s)	<input type="checkbox"/>					
Father's native language(s)	<input type="checkbox"/>					

You selected "Other" at least once on the previous question. Please name the other language(s) and to which person(s) it applies in the space below:

Which of these languages is/are spoken consistently among your immediate family members? (Select all that apply)

- English
- Spanish
- Portuguese
- French
- Italian
- Other (Please type the language(s) here) _____

Which of the following Spanish classes are you CURRENTLY taking? (Select all that apply)

- S200 Intermediate Spanish I
- S250 Intermediate Spanish II
- S280 Spanish Grammar in Context
- S308 Composition and Conversation in Spanish
- S315 Spanish in the Business World
- S317 Spanish Conversation and Diction
- S324 Intro to Hispanic Cultures
- S326 Intro to Hispanic Linguistics
- S328 Intro to Hispanic Literatures
- S334 Panoramas of Hispanic Lit
- S411 Spain: The Cultural Context
- S412 Spanish America: The Cultural Context
- S417 Hispanic Poetry
- S425 Spanish Phonetics
- S427 The Structure of Spanish
- S429 Spanish Sociolinguistics and Pragmatics
- S430 The Acquisition of Spanish
- S450 Don Quixote
- S473 Hispanic Literature and Literary Theory
- S474 Hispanic Literature and Society
- S479 Mexican Literature

Which of these Spanish classes did you take at IU anytime BEFORE this semester? (Select all that apply)

- S200 Intermediate Spanish I
- S250 Intermediate Spanish II
- S280 Spanish Grammar in Context
- S308 Composition and Conversation in Spanish
- S315 Spanish in the Business World
- S317 Spanish Conversation and Diction
- S324 Intro to Hispanic Cultures
- S326 Intro to Hispanic Linguistics
- S328 Intro to Hispanic Literatures
- S334 Panoramas of Hispanic Lit
- S407 Survey of Spanish Literature 1
- S408 Survey of Spanish Literature 2
- S411 Spain: The Cultural Context
- S412 Spanish America: The Cultural Context
- S413 Hispanic Culture in the US
- S417 Hispanic Poetry
- S418 Hispanic Drama
- S419 Modern Spanish Prose Fiction
- S422 Hispanic Cinema
- S423 The Craft of Translation
- S425 Spanish Phonetics
- S427 The Structure of Spanish
- S429 Spanish Sociolinguistics and Pragmatics
- S430 The Acquisition of Spanish
- S435 Literatura Chicana y Puertorriquena
- S450 Don Quixote
- S470 Women and Hispanic Literature
- S471 Spanish American Literature 1
- S472 Spanish American Literature 2
- S473 Hispanic Literature and Literary Theory
- S474 Hispanic Literature and Society
- S479 Mexican Literature
- S480 Argentine Literature

Which of the following best describes your major/minor at IU?

- Spanish major
- Spanish minor
- Neither

What is your Spanish major track?

- Hispanic Literature
- Hispanic Linguistics
- Hispanic Studies

What is/are your major(s) and minor(s)?

How many years have you studied Spanish in the following levels of education?

	N/A	Less than 1 year	1 to 2 years	3 or more years
Elementary School	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Middle/Junior High School	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
High School	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Select the type of Spanish courses which you have completed at each level of university study (mark all that apply). Note: A Spanish linguistics course is any course that has the purpose of discussing the scientific study of Spanish as a language and its parts (speech sounds, word structure, sentence structure, etc.). It is not the same as a grammar course.

	Grammar	Conversation	Composition (focus on writing)	Literature	Linguistics	Culture	Other
First year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Second year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Third year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fourth year course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Graduate level course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Have you ever attended a school where Spanish was the primary language of instruction? (e.g. Spanish immersion school) If yes, please explain when and for how long.

- Yes _____
- No

Have you ever visited/lived in a Spanish-speaking country for 3 weeks or more at one time?

- Yes
- No

Please provide the name of each Spanish-speaking country and city/cities you have visited/lived in for 3 weeks or more along with: (1) the date(s) you were there (e.g. 2005-2006 if full year or July 2008-September 2008) (2) the duration of the stay in months and weeks (e.g. 3 months and 2 weeks) (3) your primary purpose(s) for living there (e.g. language study, vacation, service, Peace Corps, military, mission etc.) If you have been to the same place multiple times for 3 weeks or longer, use separate lines for each time.

	Date(s)	Duration	Purpose(s)
Country and city/cities #1			
Country and city/cities #2			
Country and city/cities #3			
Country and city/cities #4			
Country and city/cities #5			
Country and city/cities #6			
Country and city/cities #7			
Country and city/cities #8			

In the last year, on average, how often did you participate outside of class in the following activities in Spanish?

	A few times or less per year	Once a month	A few times per month	Once per month	A few times or more per month
Speaking Spanish with native speakers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking Spanish with non-native speakers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Watching TV, videos, or movies in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading articles or books in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Listening to music in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Communicating online in Spanish (Facebook, Twitter, email etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texting in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have had any relatively frequent contact with native Spanish speakers within the past 2 years, please complete the following information about them and about your contact with them: Under the far left-hand column ("Native speaker #") put the sex (M/F), age (estimate if unknown) of each person (e.g., M, 24)

	Country of origin	City of origin (if known)	Approximate frequency of contact with this person in Spanish, either listening to them or speaking to them (daily, weekly, monthly, annually)	Relationship to you (family, friend, co-worker, teacher, acquaintance)
Native speaker #1				
Native speaker #2				
Native speaker #3				
Native speaker #4				
Native speaker #5				
Native speaker #6				

In your opinion, how do your skills in Spanish compare to the other second-language Spanish learners in your Spanish classes?

	Below Average	Average	Above Average	Well Above Average
Understanding spoken speech	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking (speed of speech, hesitations, pauses, & overall errors)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pronunciation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grammar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Have you studied/learned any other languages other than Spanish?

- Yes
- No

Please list the language(s) on the left and select your approximate overall proficiency in each language on the right.

Language	Beginner	Intermediate	Advanced	Near-native
Language 1	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native
Language 2	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native
Language 3	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native
Language 4	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near native
Language 5	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near native

If you were to study abroad sometime in the future in a Spanish-speaking country, please rank what your reasons for doing so would be in order of importance. Note: Each number can only have one reason checked, so they must be ranked 1 through 9. 1 = most important 9 = least important

- _____ Experience a foreign culture
- _____ Expand my circle of friends
- _____ Explore my heritage
- _____ Prepare for a future career in which I would use Spanish
- _____ Fulfill academic requirements
- _____ Improve my Spanish
- _____ See the world/travel around
- _____ Build my résumé to be more competitive in the job market
- _____ I didn't have a choice (someone else wanted me to do it)

If you were to study abroad sometime in the future in a Spanish-speaking country, please rank the following activities in terms of how important they would be to you during a study abroad experience. Note: Each number can only have one reason checked, so they must be ranked 1 through 8. 1 = most important, 8 = least important

- Studying/School work
- Making friends with native Spanish speakers from that location and spending time with them
- Local sightseeing around the location of my study abroad program
- Visiting other places outside of the location of my study abroad program
- Making friends with English-speaking students from my program and spending time with them
- Trying new cultural experiences (e.g., new foods, new activities)
- Participating in some sort of service learning in the local community
- Partying and/or Night scene

Please rank the following language learning priorities in the order of how important they are to you this semester as you study Spanish. Note: Each number can only have one reason checked, so they must be ranked 1 through 7. 1 = most important, 7 = least important

- Learn how to speak with a certain Spanish accent
- Improve my overall pronunciation
- Improve my overall grammar
- Learn how to read better in Spanish
- To better understand native speakers when they speak
- Learn how to write better in Spanish
- To better understand the cultural aspects (e.g., appropriateness, politeness etc.) of the Spanish language

Please use the slider to rate how strongly you agree or disagree with the following statements. Moving the slider farther to the right indicates a stronger agreement and farther to the left indicates a stronger disagreement. The middle (50) is neutral. Please answer honestly. Your answers will not be shared with anyone else with your name attached to them.

- _____ I love the Spanish language
- _____ I mainly study Spanish to build my résumé
- _____ Learning Spanish is boring to me
- _____ Language learning is a fun and interesting process
- _____ I am studying Spanish primarily to fulfill a language requirement at my university
- _____ I am studying Spanish primarily to improve my Spanish language skills and cultural knowledge
- _____ For me, studying Spanish is fun
- _____ I am excited to take Spanish classes this semester
- _____ I want to learn how to have pronunciation like native Spanish speakers from a certain dialect
- _____ I am not interested in learning to speak a new dialect of Spanish
- _____ I care a lot about how "native-like" my Spanish sounds
- _____ I care more about getting my point across than how my Spanish sounds

Have you had any history of trouble with hearing?

- Yes
- No

Appendix I: Language Contact and Use Questionnaire (i.e., Time 2 questionnaire – study abroad groups)

Please complete the following survey about your semester in Spain as honestly and accurately as possible.

Please enter your participant code accurately.

What is your email address?

What is the name of your home university in the U.S.?

Through what organization is your study abroad program in Seville offered?

- [Study Abroad 1] (to maintain the anonymity of the programs, the names were removed)
- [Study Abroad 2]

Which of the following programs were you a part of this semester in Seville?

- Teaching Development Program
- Liberal Arts
- Advanced Liberal Arts
- Language and Society
- Communication, New media, and Journalism

What level of Spanish courses did you take this semester (Select all that apply)

- Beginner
- Intermediate
- Advanced
- I only took courses taught in English

Please select the living arrangement you have had during your semester abroad:

- Dormitory with English-speaking roommates (or speakers of other languages apart from Spanish)
- Dormitory with Spanish-speaking roommates
- Dormitory with my own room
- Home stay with a Spanish-speaking person or family (even if another English-speaking student also lived with that same person/family)
- Apartment with other English-speaking students (or speakers of other languages apart from Spanish)
- Apartment with native Spanish-speakers
- My own apartment
- Other: Please describe _____

For each of the people with whom you lived in the same residence during this semester, please fill out the following information in the spaces provided: (In the far left column where you see "Person #1" etc, put the status of that person (host mom, host dad, host child, roommate))

	Where was this person from (city/country) originally?	Native language (English, Spanish, or write another language if applicable)	Male or Female?	Age (estimate if unknown)
Person #1				
Person #2				
Person #3				
Person #4				
Person #5				
Person #6				

For each of the people you mentioned in the previous question, how often on average did you do speak Spanish and English with them for 15-20 minutes or more at a time?

	Speak Spanish						Speak English					
	Never	Rarely	Once a week	A few times a week	Once a day	Multiple times a day	Never	Rarely	Once a week	A few times a week	Once a day	Multiple times a day
Person #1	<input type="radio"/>											
Person #2	<input type="radio"/>											
Person #3	<input type="radio"/>											
Person #4	<input type="radio"/>											
Person #5	<input type="radio"/>											
Person #6	<input type="radio"/>											

Please list all academic courses you have taken this semester with the following information:

	Taught in Spanish? (Type yes or no)	Hours per week spent in this class	Professor's name (if known)	Professor's city of origin (if known)
Course #1 name				
Course #2 name				
Course #3 name				
Course #4 name				
Course #5 name				

On an average day, how much time did you spend doing the following official academic activities? Note that dedicated listening/speaking practice is not the same as having a regular conversation with someone. Dedicated practice would be something like doing a specific listening activity or a speaking practice activity, perhaps with a partner, or an intercambio with a Spanish speaker etc...

	None	Less than 30 minutes	30 minutes to 1 hour	1-2 hours	2+ hours
Reading/studying for school in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Writing for school in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dedicated listening practice in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dedicated speaking practice in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How frequently, on average, did you communicate in the following ways in English with people back home in the U.S. during your semester abroad?

	Not at all	A few times total	Once a month	A few times a month	A few times a week	Daily
Texting	<input type="radio"/>					
Email	<input type="radio"/>					
Social media (Facebook, Twitter etc.)	<input type="radio"/>					
Handwritten/Typed letters	<input type="radio"/>					
Skype/phone calls or something similar	<input type="radio"/>					

Thinking about your overall use of Spanish versus English during the semester, please give an approximate percentage of how much SPANISH you used in general (including all language-related activities like speaking, writing, reading etc.). 100% is only Spanish and 0% is only English. (If you wish to keep the slider at 0, you must at least click it to activate it). IMPORTANT: This is not just referring to class, but to your entire life during this semester.

_____ % use of Spanish

For the following language-related activities during the semester, please use the slider to mark the approximate percentage of time spent doing that activity IN SPANISH. (If you wish to keep the slider at 0, you must at least click it to activate it). 100% is only Spanish and 0% is only English.

_____ Speaking (to anyone)

_____ Reading (includes books, internet, magazines, newspapers etc.)

_____ Writing (includes any writing such as emails, homework, social media, texting etc.)

Imagining that your average day (while awake) lasted 16 hours (7am-11pm), approximately how much of that time did you spend doing the following activities. Estimate the best you can and try to be realistic. Note that all sliders go from 0 to 16 hours in increments of 0.1 hour (6 min.) and should add up to no more than 16 hours with the numbers down the right hand side of the sliders. An increment of 0.5 = 1/2 hour. If you want to leave a slider at 0, you must at least click it to activate it.

_____ Conversing with native Spanish speakers in Spanish

_____ Conversing with native Spanish speakers in English

_____ Conversing with native English speakers in Spanish

_____ Conversing with native English speakers in English

_____ Listening to lectures in Spanish

_____ Listening to lectures in English

_____ Watching TV, movies, and/or videos online in Spanish

_____ Watching TV, movies, and/or videos online in English

_____ Reading books, study materials, newspapers, online articles and/or magazines in Spanish

_____ Reading books, study materials, newspapers, online articles, and/or magazines in English

_____ All other activities that do not involve language (e.g., dressing, walking, shopping, napping etc).

Use this to make up the rest of the 16 hours.

NOTE: All of the questions on this page incorporated a sliding scale that did not translate into Microsoft Word format.

For the following Spanish language skills, how would you compare your current abilities to your abilities at the beginning of this semester?

	Somewhat worse	About the same	Somewhat better	Much better
Understanding native speaker speech	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking (speed of speech, hesitations/pauses, errors etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Writing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pronunciation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grammar (knowledge, accuracy)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please mark the number of native Spanish speakers from Seville with whom you had actual conversations IN SPANISH (of approximately 15 to 20 minutes or more) on at least a WEEKLY basis (on average).

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12

Please mark the number of native Spanish speakers from Seville with whom you had actual conversations IN SPANISH (of approximately 15 to 20 minutes or more) on a DAILY basis (on average).

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12

Not including people with whom you lived in the same residence, list up to 12 other people with whom you had extended conversations in SPANISH (15-20 minutes or more) as well as the frequency of these conversations during your study abroad. In the space provided, please provide the speaker's sex (M or F), age (estimate if unknown), and their relationship to you (e.g. friend, classmate, program director etc.) Example: M, 23, friend. Click the box in the 'Native Spanish speaker' column if the person is a native speaker of Spanish

	Native Spanish speaker	Rarely	Once a week	Multiple times a week	Once a day	Multiple times a day
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					

Not including people with whom you lived in the same residence, list up to 12 other people with whom you had extended conversations in ENGLISH (15-20 minutes or more) as well as the frequency of these conversations during your study abroad. In the space provided, please provide the speaker's sex (M or F), age (estimate if unknown), and their relationship to you (e.g. friend, classmate, program director, etc.). Example: M, 23, friend Click the box in the 'Native Spanish speaker' column if the person is a native speaker of Spanish

	Native Spanish speaker	Rarely	Once a week	Multiple times a week	Once a day	Multiple times a day
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					

Please rank the following activities in order of how much of a priority they actually ended up being to you during this semester: 1 = most important, 8 = least important

- _____ Studying/School work
- _____ Making friends with native Spanish speakers and spending time with them
- _____ Local sightseeing around Seville
- _____ Visiting other places outside of Seville
- _____ Making friends with English-speaking students from my program and spending time with them
- _____ Trying new cultural experiences (e.g., new foods, new activities)
- _____ Participating in some sort of service learning in the local community
- _____ Partying and/or night scene

Please rank the following language learning priorities in the order of how important they actually ended up being to you this semester as you tried to improve your Spanish. In other words, what did you end up caring about the most this semester in terms of your language learning? 1 = most important, 7 = least important

- _____ Learn how to speak with the local accent
- _____ Improve my overall pronunciation
- _____ Improve my overall grammar
- _____ Learn how to read better in Spanish
- _____ To better understand native speakers when they speak
- _____ Learn how to write better in Spanish
- _____ To better understand the cultural aspects (e.g., appropriateness, politeness etc.) of the Spanish language

Please click and drag the bar to show how much you agree or disagree with the following statements. The farther right you go shows stronger agreement and the farther left shows stronger disagreement. The middle (50) is neutral. If you want to leave a slider in the middle, you must at least click it to activate it. Please answer honestly. Your answers will not be shared with anyone with your name attached to them. **A sliding scale was used, but did not translate to Word format.**

- I like the people from Seville
- I wish the lifestyle at home in the U.S. was more like the lifestyle in Seville
- I dislike the culture in Seville
- I wish I had studied abroad in a different place than Seville.
- I love the Spanish language
- I mainly study Spanish to build my résumé
- Learning Spanish is boring to me
- Language learning is a fun and interesting process
- I think the people of Seville speak Spanish poorly
- I think the Spanish spoken in Seville is neat
- I dislike the pronunciation of words in the Spanish of Seville
- If I lived in Seville longer, I would try to speak like the people here
- I mainly studied abroad to fulfill a language requirement at my university
- I mainly studied abroad to improve my Spanish language skills and experience the language in its cultural context
- For me, studying abroad in Seville is more about having fun than about learning Spanish
- I am glad that I have had to take Spanish classes while abroad.
- When I speak Spanish I consciously try to sound like a local from Seville
- I care more about getting my point across than about how I sound while speaking
- I have my own accent when speaking and avoid speaking like the locals in Seville
- I care a lot about how native-like my Spanish sounds

During your time in Seville, did you ever notice or learn anything unique about how people from Seville tend to pronounce S in words such as *estás*, *fiesta*, *esto*, *España*, or *esquina*?

- Yes
- No

Please describe what you noticed and/or learned about how S is pronounced in words such as the ones above: *estás*, *fiesta*, *esto*, *España*, *esquina*. What does it sound like?

How did you learn of this way of pronouncing S in Seville? (Select all that apply)

- I noticed it here when people were speaking
- Someone told me about it before coming to Seville
- Someone told me about it after coming to Seville
- I read about it somewhere before coming to Seville
- I read about it somewhere after coming to Seville
- TV or movies
- We talked about it in a class I had before studying abroad in Seville
- We talked about it in a class I have had this semester in Seville

Had you ever met anyone from southern Spain (e.g., Seville, Granada, Málaga, etc) before coming to Seville this semester and heard them speak Spanish?

- No
- Yes, I had a friend from southern Spain before studying here
- Yes, I had a teacher from southern Spain before studying here
- Yes, I had a family member from southern Spain before studying here
- Yes, other (please explain) _____

Had you ever met anyone from Spanish speaking countries in the Caribbean before coming to Seville this semester?

- No
- Yes, I had a friend from the Caribbean before studying here
- Yes, I had a teacher from the Caribbean before studying here
- Yes, I had a family member from the Caribbean before studying here
- Yes, other (please explain) _____

Do you think your pronunciation has become more like the native speakers from Seville? In other words, have you tried to speak like them in any way or do you avoid speaking like native speakers from Seville? Please explain.

Appendix J: Language Contact and Use Questionnaire (Time 2 – At-home group)

Please answer the questions in this survey about this fall semester at IU as honestly and accurately as possible.

Please enter your participant code accurately.

What is your IU email address?

Please list all Spanish courses you have been taking this semester with the following information:

	Taught in Spanish? (Type yes or no)	Hours per week spent in this class	Professor's name	Professor's place of origin (if known)
Course #1 name				
Course #2 name				
Course #3 name				
Course #4 name				
Course #5 name				

On an average day, how much time did you spend doing the following official academic activities? Note that dedicated listening/speaking practice is not the same as having a normal conversation with someone. Dedicated practice would be something like doing a specific listening activity or a speaking practice activity, perhaps with a partner etc...

	None	Less than 30 minutes	30 minutes to 1 hour	1-2 hours	2+ hours
Reading/studying for school in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Writing for school in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dedicated listening practice in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dedicated speaking practice in Spanish	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How frequently, on average, did you communicate in the following ways in Spanish with people during this semester?

	Not at all	A few times total	Once a month	A few times a month	A few times a week	Daily
Texting	<input type="radio"/>					
Email	<input type="radio"/>					
Social media (Facebook, Twitter etc.)	<input type="radio"/>					
Handwritten/Typed letters	<input type="radio"/>					
Phone/Skype	<input type="radio"/>					
In person (speaking)	<input type="radio"/>					

How frequently, on average, did you communicate in the following ways in English with people during this semester?

	Not at all	A few times total	Once a month	A few times a month	A few times a week	Daily
Texting	<input type="radio"/>					
Email	<input type="radio"/>					
Social media (Facebook, Twitter etc.)	<input type="radio"/>					
Handwritten/Typed letters	<input type="radio"/>					
Phone/Skype calls	<input type="radio"/>					
In person (speaking)	<input type="radio"/>					

Thinking about your overall use of Spanish during the semester, please give an approximate percentage of how much SPANISH you used in general compared to English (including all language-related activities like speaking, writing, reading etc.). 100% means that you used only Spanish the entire semester for communication and 0% means you only used English and not any Spanish for any communications. IMPORTANT: This is not just referring to class, but your entire life during the semester.

_____ % use of Spanish

Note: These questions and the previous question were presented using a sliding scale in Qualtrics, which did not translate well to Microsoft Word format.

For the following activities during the entire semester, please use the slider to mark the approximate percentage of time spent doing that activity IN SPANISH compared to English. (If you wish to keep the slider at 0, you must at least click it to activate it). 100% is only Spanish and 0% is only English.

_____ Speaking (to anyone)

_____ Reading (includes books, internet, magazines, newspapers etc.)

_____ Writing (includes any writing such as emails, homework, social media, texting etc.)

Imagining that your average day (while awake) lasted 16 hours (7am-11pm), approximately how much of that time did you spend doing the following activities. Estimate the best you can and try to be realistic. Note that all sliders go from 0 to 16 hours in increments of 0.1 hour (6 min.) and should add up to no more than 16 hours with the numbers down the right hand side of the sliders. An increment of 0.5 = 1/2 hour. If you want to leave a slider at 0, you must at least click it to activate it.

_____ Conversing with native Spanish speakers in Spanish

_____ Conversing with native Spanish speakers in English

_____ Conversing with native English speakers in Spanish

_____ Conversing with native English speakers in English

_____ Listening to lectures in Spanish

_____ Listening to lectures in English

_____ Watching TV, movies, and/or videos online in Spanish

_____ Watching TV, movies, and/or videos online in English

_____ Reading books, study materials, newspapers, online articles and/or magazines in Spanish

_____ Reading books, study materials, newspapers, online articles, and/or magazines in English

_____ All other activities that do not involve language (e.g., dressing, walking, shopping, napping etc)

For the following Spanish language skills, how would you compare your current abilities to your abilities at the beginning of this semester?

	Somewhat worse	About the same	Somewhat better	Much better
Understanding native speaker speech	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking (speed of speech, hesitations/pauses, errors etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Writing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pronunciation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grammar (knowledge, accuracy)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please mark the number of native Spanish speakers with whom you had actual conversations IN SPANISH during this fall semester (about 15 to 20 minutes or more) on at least a WEEKLY basis (on average).

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12

Please mark the number of native Spanish speakers with whom you had actual conversations IN SPANISH during this fall semester (about 15 to 20 minutes or more) on a DAILY basis (on average).

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12

List up to 12 other people with whom you had extended conversations in SPANISH (15-20 minutes or more) as well as the frequency of these conversations during this semester. In the space provided, please provide the speaker's sex (M or F), age (estimate if unknown), and their relationship to you (e.g. friend, classmate, instructor etc.) Example: M, 45, instructor Click the box in the 'Native Spanish speaker' column if the person is a native speaker of Spanish

	Native Spanish speaker	Rarely	Once a week	Multiple times a week	Once a day	Multiple times a day
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					

List up to 12 other people with whom you had extended conversations in ENGLISH (15-20 minutes or more) as well as the frequency of these conversations during this semester. In the space provided, please provide the speaker's sex (M or F), age (estimate if unknown), and their relationship to you (e.g. friend, classmate, instructor, etc.). Example: M, 22, friend. Click the box in the 'Native Spanish speaker' column if the person is a native speaker of Spanish.

	Native Spanish speaker	Rarely	Once a week	Multiple times a week	Once a day	Multiple times a day
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					
sex, age, relationship	<input type="checkbox"/>					

Please rank the following language learning priorities in the order of how important they actually ended up being to you this semester as you tried to improve your Spanish (note: only one priority can be chosen for each number 1-7): 1 = most important, 7 = least important

- _____ Learn how to speak with the a certain Spanish accent
- _____ Improve my overall pronunciation
- _____ Improve my overall grammar
- _____ Learn how to read better in Spanish
- _____ To better understand native speakers when they speak
- _____ Learn how to write better in Spanish
- _____ To better understand the cultural aspects (e.g., appropriateness, politeness etc.) of the Spanish language

Please click and drag the bar to show how much you agree or disagree with the following statements. The farther right you go shows stronger agreement and the farther left shows stronger disagreement. The middle (50) is neutral. If you want to leave a slider in the middle, you must at least click it to activate it. Please answer honestly. Your answers will not be shared with anyone with your name attached to them.

- _____ I like people who speak Spanish
- _____ I wish the lifestyle at home in the U.S. was more like the lifestyle in Spanish-speaking countries
- _____ I dislike Hispanic culture
- _____ I would like to study abroad in a Spanish-speaking country someday
- _____ I love the Spanish language
- _____ I mainly study Spanish to build my résumé
- _____ Learning Spanish is boring to me
- _____ Language learning is a fun and interesting process
- _____ I think the people of certain dialects speak Spanish poorly
- _____ I think the Spanish spoken in Spain is neat
- _____ I dislike the pronunciation of words in the Spanish of Spain
- _____ If I lived in a Spanish-speaking country for a while (like studying abroad), I would try to speak like the people there
- _____ I am mainly in Spanish class to fulfill a language requirement at IU
- _____ I am mainly taking Spanish to improve my Spanish language skills and learn more about Hispanic culture.
- _____ If I were to study abroad in a Spanish-speaking country, it would be more about having fun than about learning Spanish
- _____ I am glad that I have had to take Spanish classes this semester.
- _____ When I speak Spanish I consciously try to sound like a native speaker from a certain place in the Spanish-speaking world
- _____ I care more about getting my point across than about how I sound while speaking Spanish
- _____ I have my own accent when speaking and avoid speaking like native Spanish speakers

_____ I care a lot about how native-like my Spanish sounds

Did you have any contact during this semester with native Spanish speakers from parts of southern Spain, such as Seville, Córdoba, Granada, Cádiz, Málaga, or from Spanish-speaking countries in the Caribbean?

- Yes
- No

Please give the following information about the people with whom you have had contact who are from southern Spain or the Caribbean Under the far left column, in the blank space under "sex, age, city, relationship," put the following information in the format (M, 25, Puerto Rico, friend). Sex - M or F (male or female). Age - Estimate if unknown City - if you know the city the person is from Relationship to you - teacher, friend, host mom etc... Then, mark the frequency of contact that you had with that person in terms of having conversations with them and/or listening to them speak.

	None/Rarely	A few times a month	Once a week	2-4 times a week	Daily
sex, age, city, relationship	<input type="radio"/>				
sex, age, city, relationship	<input type="radio"/>				
sex, age, city, relationship	<input type="radio"/>				
sex, age, city, relationship	<input type="radio"/>				
sex, age, city, relationship	<input type="radio"/>				
sex, age, city, relationship	<input type="radio"/>				

Even though you have not had contact with Spanish speakers from southern Spain or the Caribbean, were you made aware, through any means such as television, books, the internet, class, or conversations with other people, of any peculiarities of how people in southern Spain or the Caribbean pronounce words?

- Yes
- No

Please describe in the box below anything that you heard about peculiarities of pronunciation in the Spanish of southern Spain and how you heard about it/them.

Let's get a little more specific. Have you ever heard anything about how people from southern Spain or the Caribbean tend to pronounce S in words such as "estás, fiesta, or esto"

- Yes
- No

How did you hear about this way of pronouncing S in southern Spain (i.e., Andalusia) or the Caribbean?
(Select all that apply)

- I noticed it when people were speaking
- Someone told me about it
- I read about it somewhere (If so, where?) _____
- TV or movies
- We talked about it in a class I had before this semester
- We talked about it in a class I have had this semester

Have you ever heard of something called "S weakening" or "S aspiration" in Spanish?

- Yes
- No

Please describe briefly as best you can what S weakening/aspiration is.

Appendix K: Native Speaker Questionnaire

¿Cuál es el código de participante que se le ha asignado? (Asegúrese de que sea correcto al deletrear)

¿Es usted...

- hombre?
- mujer?

¿Cuántos años tiene usted?

¿De dónde es usted originalmente? (país y ciudad)

¿Cuál de las siguientes oraciones mejor describe la cantidad de tiempo que usted lleva en el lugar de origen que acaba de mencionar en la pregunta anterior?

- Llevo toda la vida en mi ciudad de origen.
- Llevo la mayoría de la vida en mi ciudad de origen pero pasé una temporada viviendo en otro(s) lugar(es).
- Llevo menos de la mitad de la vida en mi ciudad de origen porque viví en otro(s) lugar(es) por mucho tiempo.
- Llevo la mayoría de la vida en un lugar diferente de mi ciudad de origen pero he vivido por una temporada en mi lugar de origen.
- Llevo toda la vida en un lugar diferente de mi ciudad de origen.

¿En qué otros lugares ha vivido usted y por cuánto tiempo? (escriba la ciudad y el país, y seleccione la duración apropiada)

Lugar 1	<input type="radio"/> 1 año o menos	<input type="radio"/> de 2 a 5 años	<input type="radio"/> más de 5 años	<input type="radio"/> de 6 a 10 años	<input type="radio"/> más de 10 años
Lugar 2	<input type="radio"/> 1 año o menos	<input type="radio"/> de 2 a 5 años	<input type="radio"/> más de 5 años	<input type="radio"/> de 6 a 10 años	<input type="radio"/> más de 10 años
Lugar 3	<input type="radio"/> 1 año o menos	<input type="radio"/> de 2 a 5 años	<input type="radio"/> más de 5 años	<input type="radio"/> de 6 a 10 años	<input type="radio"/> más de 10 años
Lugar 4	<input type="radio"/> 1 año o menos	<input type="radio"/> de 2 a 5 años	<input type="radio"/> más de 5 años	<input type="radio"/> de 6 a 10 años	<input type="radio"/> más de 10 años
Lugar 5	<input type="radio"/> 1 año o menos	<input type="radio"/> de 2 a 5 años	<input type="radio"/> más de 5 años	<input type="radio"/> de 6 a 10 años	<input type="radio"/> más de 10 años

¿Dónde vive usted ahora?

- Estados Unidos
- España

¿Por cuánto tiempo ha vivido usted en Sevilla?

- 1 año o menos
- de 2 a 5 años
- de 6 a 10 años
- de 10 a 15 años
- de 15 a 20 años
- más de 20 años

¿Por cuánto tiempo ha vivido usted en Estados Unidos?

- 1 año o menos
- de 2 a 5 años
- de 6 a 10 años
- de 10 a 15 años
- de 15 a 20 años
- más de 20 años

¿Habla usted otro(s) idioma(s) aparte del español?

- Sí
- No

Para los idiomas que usted sabe hablar, marque el nivel de habilidad que tiene con cada uno de ellos. Usted puede dejar en blanco los idiomas que no sabe hablar.

Inglés	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Italiano	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Portugués	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Francés	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Catalán	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Euskera	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Alemán	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Árabe	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo
Otro (escríbalo, por favor)	<input type="radio"/> Principiante	<input type="radio"/> Intermedio	<input type="radio"/> Avanzado	<input type="radio"/> Casi nativo	<input type="radio"/> Nativo

¿Cuál es el nivel más alto de educación que usted ha logrado?

- Primaria
- Secundaria obligatoria
- Bachillerato
- Soy alumno universitario actualmente
- Licenciatura
- Formación profesional
- Maestría
- Doctorado

Si usted tiene trabajo actualmente, ¿en qué trabaja? (dé una descripción o título de una o dos palabras si es posible)

English Translation of the Native Speaker Questionnaire

What is the participant code that has been assigned to you? (Please be sure that it is spelled correctly)

Are you...

- male?
- female?

How old are you?

Where are you originally from? (city and country)

Which of the following sentences best describes the quantity of time that you have spent in the place of origin that you just mentioned in the previous question?

- I've spent my whole life in my city of origin.
- I've spent the majority of my life in my city of origin but I spent time living in other places.
- I've spent less than half of my life in my city of origin because I lived in other places for a long time.
- I have spent the majority of my life in a place other than my city of origin but I lived for a time in my city of origin.
- I have spent my whole life in a place other than my city of origin.

In what other places have you lived and for how long? (write the city, country, and select the appropriate duration)

Place 1	<input type="radio"/> 1 year or less	<input type="radio"/> 2-5 years	<input type="radio"/> More than 5 years	<input type="radio"/> 6-10 years	<input type="radio"/> More than 10 years
Place 2	<input type="radio"/> 1 year or less	<input type="radio"/> 2-5 years	<input type="radio"/> More than 5 years	<input type="radio"/> 6-10 years	<input type="radio"/> More than 10 years
Place 3	<input type="radio"/> 1 year or less	<input type="radio"/> 2-5 years	<input type="radio"/> More than 5 years	<input type="radio"/> 6-10 years	<input type="radio"/> More than 10 years
Place 4	<input type="radio"/> 1 year or less	<input type="radio"/> 2-5 years	<input type="radio"/> More than 5 years	<input type="radio"/> 6-10 years	<input type="radio"/> More than 10 years
Place 5	<input type="radio"/> 1 year or less	<input type="radio"/> 2-5 years	<input type="radio"/> More than 5 years	<input type="radio"/> 6-10 years	<input type="radio"/> More than 10 years

Where do you live now?

- United States
- Spain

How long have you lived in Seville? (Only for the SEVILLE group)

- 1 year or less
- 2 to 5 years
- 6 to 10 years
- 10 to 15 years
- 15 to 20 years
- More than 20 years

How long have you lived in the United States?

- 1 year or less
- 2 to 5 years
- 6 to 10 years
- 10 to 15 years
- 15 to 20 years
- More than 20 years

Do you speak any other languages?

- Yes
- No

For the languages that you know how to speak, mark the skill level that you have for each one. You can leave blank the languages that you do not know how to speak.

English	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
Italian	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
Portuguese	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
French	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
Catalan	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
Basque	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
German	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
Arabic	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native
Other (please write it)	<input type="radio"/> Beginner	<input type="radio"/> Intermediate	<input type="radio"/> Advanced	<input type="radio"/> Near-native	<input type="radio"/> Native

What is the highest level of education that you have achieved?

- Elementary
- Obligatory secondary
- High school
- I am currently a university student
- Bachelor's degree
- Professional training
- Master's degree
- Doctorate

If you currently have a job, what is it? (Give a description or title of one or two words if possible).

Appendix L: Word Familiarity Questionnaire

Please read the following list of words and select from the three answer options which applies to your level of familiarity with that word. If you know the word's meaning, type a one word English translation in the box provided.

Please type your participant code accurately

gusto

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

éste

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

busca

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

usted

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

puesto

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

presta

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

respeto

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

oscuro

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

esposa

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

asiste

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

discurso

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

estilo

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

revista

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

honesto

- I know what this word means (please write what it means in English in the space below)

- I have heard/seen this word before, but I do not know its meaning
 I have never heard/seen this word before and do not know its meaning

responde

- I know what this word means (please write what it means in English in the space below)

- I have heard/seen this word before, but I do not know its meaning
 I have never heard/seen this word before and do not know its meaning

esto

- I know what this word means (please write what it means in English in the space below)

- I have heard/seen this word before, but I do not know its meaning
 I have never heard/seen this word before and do not know its meaning

fiesta

- I know what this word means (please write what it means in English in the space below)

- I have heard/seen this word before, but I do not know its meaning
 I have never heard/seen this word before and do not know its meaning

justo

- I know what this word means (please write what it means in English in the space below)

- I have heard/seen this word before, but I do not know its meaning
 I have never heard/seen this word before and do not know its meaning

hispano

- I know what this word means (please write what it means in English in the space below)

- I have heard/seen this word before, but I do not know its meaning
 I have never heard/seen this word before and do not know its meaning

escucha

- I know what this word means (please write what it means in English in the space below)

- I have heard/seen this word before, but I do not know its meaning
 I have never heard/seen this word before and do not know its meaning

espera

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

disco

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

bosque

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

amistad

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

artista

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

triste

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

oeste

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

espacio

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

supuesto

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

esta

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

destino

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

aspecto

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

discute

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

fresco

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

escribir

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

descanso

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

aspira

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

opuesto

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

espejo

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

estado

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

escapa

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

sistema

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

después

I know what this word means (please write what it means in English in the space below)

I have heard/seen this word before, but I do not know its meaning

I have never heard/seen this word before and do not know its meaning

Appendix M: Scatterplots of correlations between language use factors and accuracy in the aspirated condition for both tasks

FORCED-CHOICE IDENTIFICATION

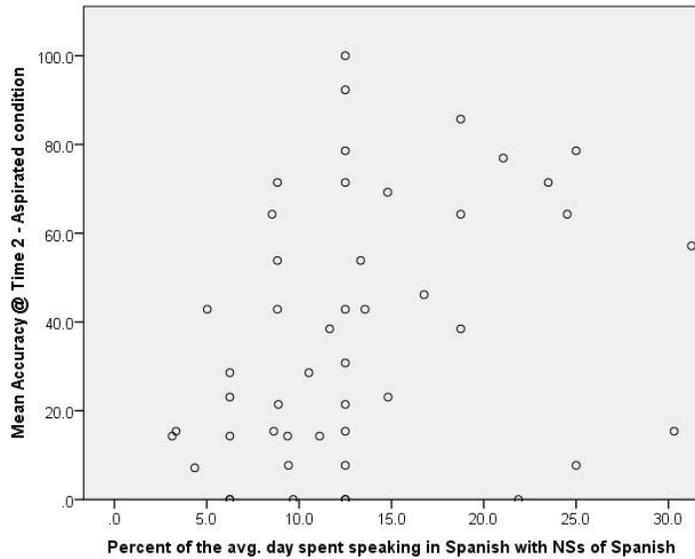


Figure 33. Correlation between mean ID accuracy at Time 2 in the aspirated condition and the reported percentage of the average day spent speaking in Spanish with NS of Spanish ($\rho = .377, p < .01$)

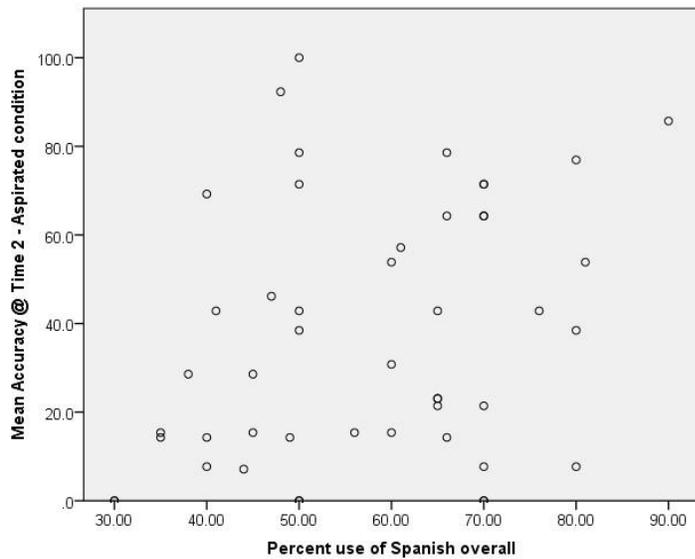


Figure 34. Correlation between mean ID accuracy at Time 2 in the aspirated condition and the reported percent use of Spanish overall ($\rho = .285, p = .05$)

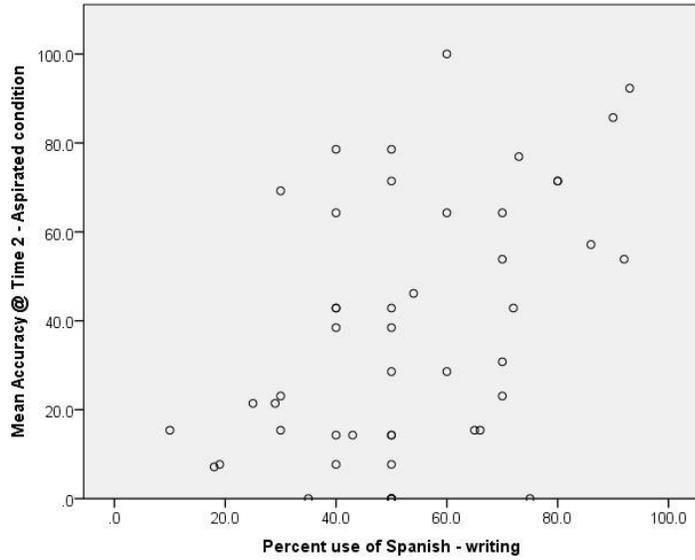


Figure 35. Correlation between mean ID accuracy at Time 2 in the aspirated condition and the reported percent use of Spanish while writing ($\rho = .417, p < .01$)

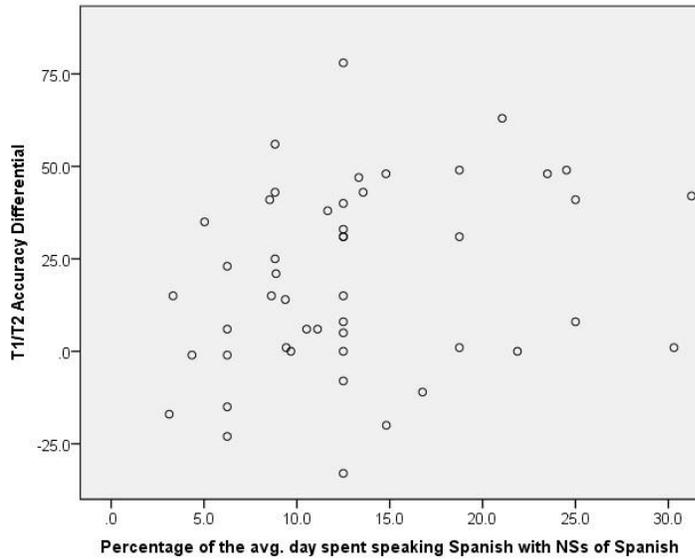


Figure 36. Correlation between T1/T2 ID accuracy differential (aspirated) and the reported percentage of the average day spent speaking Spanish with NSs of Spanish. ($\rho = .300, p < .05$)

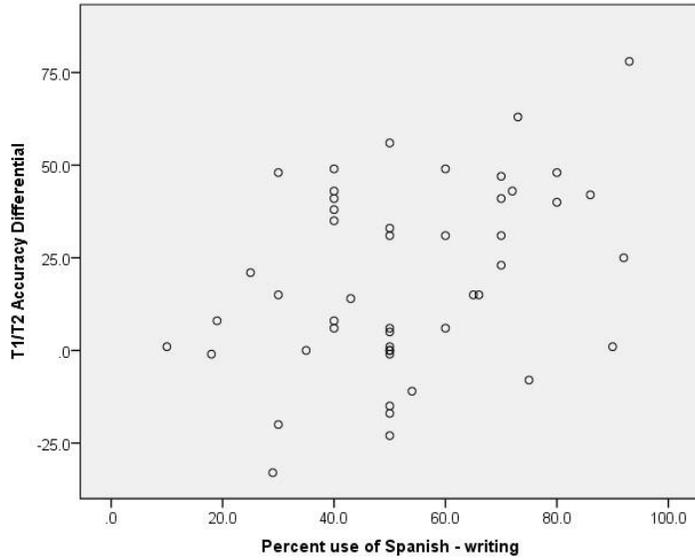


Figure 37. Correlation between T1/T2 ID accuracy differential (aspirated) and the reported percent use of Spanish while writing. ($\rho = .320, p < .05$)

LEXICAL DECISION

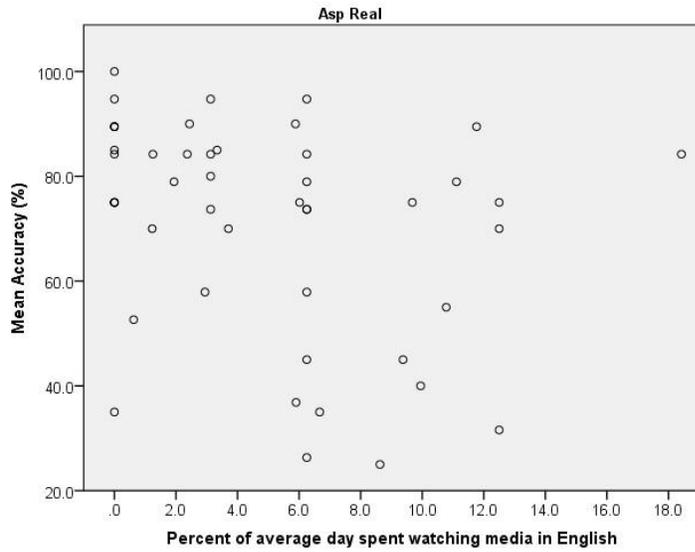


Figure 38. Correlation between mean lexical decision accuracy and the percent of an average day spent watching media in English ($\rho = -.373, p < .01$)

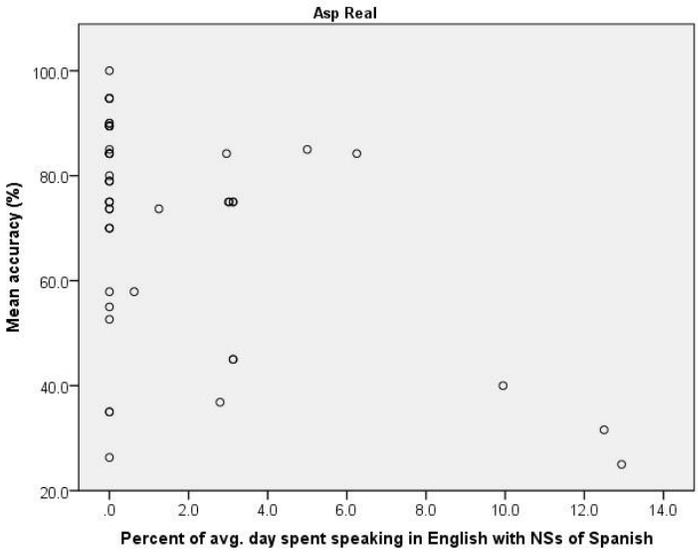


Figure 39. Correlation between mean lexical decision accuracy and the percent of an average day spent speaking in English with NSs of Spanish ($\rho = -.357, p < .05$)

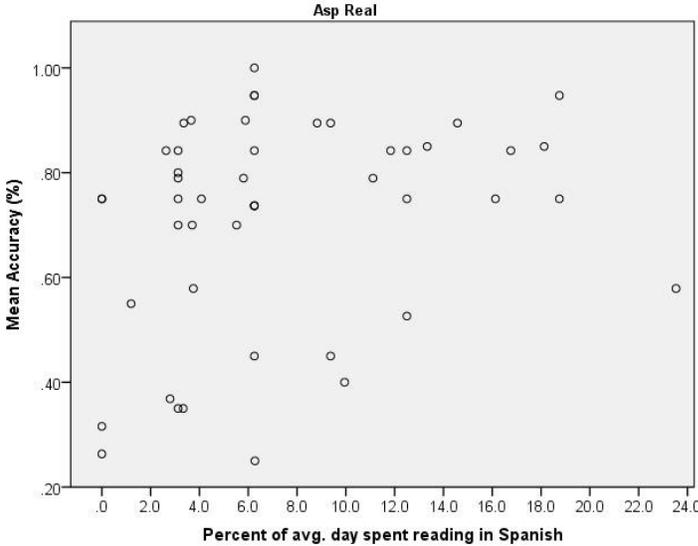


Figure 40. Correlation between mean lexical decision accuracy and the percent of an average day spent reading in Spanish ($\rho = .290, p < .05$)

Appendix N: Target language use and Contact with NS descriptives by group (Study Abroad Groups)

Table 70. Descriptive statistics for the language use factors as reported by the study abroad learners on the questionnaire at Time 2

		Factors	Mean	SD	Min	Max
Study Abroad 1 Group		% use of Spanish overall during the semester	62.2%	13.3%	30.0%	81.0%
		% use of Spanish when writing during the semester	52.3%	21.3%	10.0%	92.0%
		% use of Spanish when reading during the semester	53.0%	20.9%	10.0%	91.0%
		% use of Spanish when speaking during the semester	60.0%	15.2%	24.0%	90.0%
		% of avg. day speaking Spanish with NS of Spanish	15.7%	8.1%	3.0%	31.0%
		% of avg. day speaking Spanish with NS of English	12.2%	7.4%	3.0%	27.0%
		% of avg. day speaking English with NS of Spanish	1.0%	1.7%	0.0%	6.0%
		% of avg. day speaking English with NS of English	14.4%	6.6%	3.0%	29.0%
		% of avg. day spent watching media in Spanish	4.5%	6.8%	0.0%	31.0%
		% of avg. day spent watching media in English	4.4%	4.1%	0.0%	13.0%
		% of avg. day spent reading in Spanish	6.2%	5.2%	0.0%	19.0%
		% of avg. day spent reading in English	5.6%	4.1%	0.0%	16.0%
		Factors	Mean	SD	Min	Max
Study Abroad 2 Group		% use of Spanish overall during the semester	53.1%	15.2%	30.0%	90.0%
		% use of Spanish when writing during the semester	53.3%	19.3%	18.0%	93.0%
		% use of Spanish when reading during the semester	54.6%	25.7%	15.0%	91.0%
		% use of Spanish speaking	52.9%	17.5%	30.0%	90.0%
		% of avg. day speaking Spanish with NS of Spanish	10.8%	4.2%	3.0%	21.0%
		% of avg. day speaking Spanish with NS of English	6.7%	6.4%	0.0%	26.0%
		% of avg. day speaking English with NS of Spanish	2.0%	3.9%	0.0%	13.0%
		% of avg. day speaking English with NS of English	15.3%	8.3%	1.0%	31.0%
		% of avg. day spent watching media in Spanish	4.5%	4.4%	0.0%	16.0%
		% of avg. day spent watching media in English	5.3%	4.9%	0.0%	18.0%
		% of avg. day spent reading in Spanish	8.7%	6.0%	0.0%	24.0%
		% of avg. day spent reading in English	5.2%	3.2%	0.0%	13.0%

Table 71. Descriptive statistics for the Contact with Native Speakers factors as reported by the study abroad learners on the questionnaire at Time 2 (out of 12 possible native speakers for each variable)

	Factors	Mean	SD	Min	Max
Study Abroad 1 Group	Weekly Contact (# of NSs)	5.0	2.4	2.0	11.0
	Daily Contact (# of NSs)	3.1	2.0	1.0	8.0
		Mean	SD	Min	Max
Study Abroad 2 Group	Weekly Contact (# of NSs)	4.5	2.8	1.0	12.0
	Daily Contact (# of NSs)	2.5	1.8	0.0	8.0

Robert W. Bedinghaus

Department of Spanish and Portuguese
Indiana University
1020 E. Kirkwood, Ballantine Hall 844
Bloomington, IN 47405
Email: rbedingh@indiana.edu

EDUCATION

Indiana University, Bloomington, Indiana

Ph.D., Hispanic Linguistics, June 2015

Dissertation: The effect of exposure to phonological variation on perceptual categorization and lexical access in second language Spanish: The case of /s/-aspiration in Andalusian Spanish

Committee: Dr. Kimberly Geeslin (Chair), Dr. Isabelle Darcy, Dr. Erik Willis, Dr. Manuel Díaz-Campos

Indiana University, Bloomington, Indiana

M.A., Hispanic Linguistics, May 2011

Taylor University, Upland, Indiana

B.A., Communication Studies, May 2008

B.A., Spanish, May 2008

REFEREED PUBLICATIONS

2014

Bedinghaus, Rob, & Sedó, Beatriz. (2014). Intervocalic /d/ deletion in Málaga: Frequency effects and linguistic factors. In M. Díaz-Campos, S. Bongiovanni, K. Ebarb, & V. Filimonova (Eds.), *Indiana University Linguistics Club Working Papers*, 14.

2013

Bedinghaus, Rob. (2013). ¿Vas pa(ra) Málaga? The Reduction of *para* in Málaga, Spain: Effects of Frequency, Syntactic Category, and Social Factors. In *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, ed. Jennifer Cabrelli Amaro et al., 238-252. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2937.

REFEREED CONFERENCE PRESENTATIONS

2013

Bedinghaus, Rob. Task Variation in Speech Perception: Comparing ABX and Sequence Recall as tests of segmental discrimination. Paper presented at the *Second Language Research Forum*. Brigham Young University, October, 2013.

2012

Bedinghaus, Rob. ¿Vas pa(ra) Málaga? The reduction of *para* in Málaga: Frequency effects and (socio)linguistic factors. Paper presented at the 16th *Hispanic Linguistics Symposium*. University of Florida, October 2012.

Bedinghaus, Rob, & Sedó, Beatriz. Intervocalic /d/ deletion in Malaga: Frequency effects and linguistic factors. Paper presented at *New Ways of Analyzing Variation*. Indiana University, October 2012.

Bedinghaus, Rob. Perceptual confusion in L2 Spanish: The role of stress in the identification of Spanish [t], [d] and [r] in word-final syllables. Paper presented at *Current Approaches to Spanish and Portuguese Second Language Phonology*. University of South Carolina, February 2012.

ACADEMIC GRANTS AND AWARDS

2012

Conference travel grant - Department of Spanish and Portuguese travel grant to present research at the 16th Hispanic Linguistics Symposium, October 2012. Gainesville, FL.

Conference travel grant - Department of Spanish and Portuguese travel grant to present research at Current Approaches to Second Language Spanish and Portuguese. February, 2012. Columbia, SC.

2010

Agapito Rey Summer Fellowship - Summer award granted by the Department of Spanish and Portuguese at Indiana University.

OTHER RESEARCH AND FIELDWORK EXPERIENCE

November 2014 – Seville, Spain
Dissertation data collection

September 2014 – Seville, Spain
Dissertation data collection

October 2013

Collaborated in preparing the pre-final draft for the book: Geeslin, Kimberly, & Yim Long, Avizia. (to appear). Sociolinguistics and Second Language Acquisition: Learning to use language in context. NY: Routledge.

SERVICE TO THE PROFESSION

September 2012

Manuscript reviewer for a chapter from: Geeslin, Kimberly (Ed.). (2013). The Handbook of Spanish Second Language Acquisition. Wiley-Blackwell.

TEACHING EXPERIENCE

Fall 2009-Spring 2015: Associate Instructor (AI). Indiana University, Department of Spanish and Portuguese.

Courses taught:

S326	Introduction to Hispanic Linguistics	Spring 2014
S280	Spanish Grammar in Context	Fall 2012, 2013, 2014
S250	Intermediate Spanish II	Spring 2011, 2012, 2013, 2015; Summer, 2014
S200	Intermediate Spanish I	Spring 2010; Fall 2010, 2011; Summer 2012, 2013
S105	Accelerated first-year Spanish	Fall 2009

AI Duties: Teaching. Course preparation. Test writing. Assignment preparation. Grading.

Other:

Summer 2011: Curriculum developer and teacher for two high-school-level homeschool students.

Duties: Worked with an educational consultant to develop and implement a Spanish curriculum for the students.

GUEST LECTURES IN CLASSES

2014

“The effect of exposure to phonological variation on perceptual categorization and lexical access in second language Spanish: The case of coda /s/-weakening in Andalusian Spanish. Preliminary findings.” Guest lecture given in S430 – Spanish Second Language Acquisition. Bloomington, Indiana. October 23, 2014.

“Comparing the effectiveness of two types of speech perception tasks: ABX and Sequence Recall.” Guest lecture given in S326 – Introduction to Hispanic Linguistics. Bloomington, Indiana. January 31, 2014.

2012

“An Introduction to the Role of Transfer in Second Language Spanish Phonology.” Guest lecture given in S326 – Introduction to Hispanic Linguistics. Bloomington, Indiana. April 20, 2012.

SERVICES AND COMMITTEES

2011-2012

Treasurer, Graduate Student Advisory Committee
Department of Spanish and Portuguese, Indiana University

LANGUAGES

English, Native speaker
Spanish, Near-native speaker
Portuguese, Reading proficiency
Italian, Reading proficiency
French, Reading proficiency