NORTHWESTERN UNIVERSITY

Early versus Extended Exposure in Speech and Vocabulary Learning: Evidence from Switched-dominance Bilinguals

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Linguistics

By

Michael Blasingame

EVANSTON, ILLINOIS

March 2018

ABSTRACT

Both the timing (i.e., 'when') and amount (i.e., 'how much') of language exposure have been shown to affect language-learning outcomes. Monolinguals and (most) bilinguals confound these two factors of early exposure and extended exposure (i.e., their first-acquired language is their most used or dominant language), making it difficult to isolate the benefits that either one of these exposure patterns could provide independently for language acquisition. Switched-dominance bilinguals (i.e., heritage speakers) dissociate early and extended exposure as their first-acquired language (L1) is considerably weaker (non-dominant) compared to their stronger (dominant) second-acquired language (L2). This dissociation allows us to examine the unique benefits of both early and extended exposure on language acquisition. The current study focuses on how these exposure patterns affect speech and vocabulary learning in heritage speakers (L2-dominant) in three separate experimental paradigms.

In Experiment 1, Spanish heritage speakers (SHS) recorded sentences in Spanish (their non-dominant L1) and English (their dominant L2) along with L1-dominant Spanish and English controls in their respective (dominant) L1s. These sentences, embedded in noise at two signal-to-noise ratios (-4 dB and -8 dB signal-to-noise ratio; SNR), were presented aurally to L1-dominant listeners of Spanish and English, respectively. At the easier SNR (-4 dB SNR), SHS showed no differences in intelligibility across languages with both their English and Spanish scores reaching L1-dominant control levels of speech intelligibility. At the harder SNR (-8 dB SNR), SHS English intelligibility matched that of English L1-dominant controls, yet SHS Spanish intelligibility was significantly lower compared to that of Spanish L1 controls.

In Experiments 2 and 3, Spanish heritage speakers (SHS, L2-dominant English) performed a lexical decision task (Experiment 2) and single-word reading task (Experiment 3) in

both English and Spanish along with L1-dominant English and Spanish controls, respectively. The stimuli of interest varied orthogonally on age of acquisition (AoA) and lexical frequency, two factors known to affect word retrieval. In their dominant L2 English, SHS received similar benefits (i.e., faster reaction times and/or shorter word durations) of early-acquired and high frequency words compared to L1-dominant English controls and were not more adversely affected (i.e., slower reaction times and/or longer word durations) on late-acquired or low frequency words in English. In their non-dominant L1 Spanish, SHS were slower to respond to words over all (Experiments 2 and 3) and produced longer word durations (Experiment 3) compared to L1-dominant Spanish Controls. SHS were also more adversely affected (i.e., slower reaction times in Experiment 2 and 3; longer word durations in Experiment 3) by late-acquired and low frequency Spanish words compared to L1-dominant Spanish controls.

Combined, these results suggest that the benefits of early exposure to language may be limited in some areas of speech and vocabulary processing and that extended language usage, resulting in language dominance, may be sufficient to overcome any processing difficulties incurred in the initial delay to the L2. Furthermore, these data challenge what is meant by the term “native speaker” when modeling bilingualism on an L1-L2 distinction, as such a dimension may be unable to accurately predict the linguistic performance of some bilingual speakers.

ACKNOWLEDGMENTS

I would like to start off by thanking Ann Bradlow, my advisor, for encouraging and guiding me throughout this entire journey, as well as the feedback of Matt Goldrick and Viorica Marian, my other committee members, for their insight and support. I would also like to thank Chun Liang Chan for his amazing technological support in designing the interface for all of my data collection, as well as the Speech Communication Research Group for all of their feedback in the initial design of the study, with a special thanks to Vanessa Dopker for her support in assisting with initial data collection. Other acknowledgments go to The Graduate School for their generous Graduate Research Grant to fund this project, the Linguistics Department for their continued support, plus all of the students in the International Summer Institute, English Language Program and Spanish-speaking community for their participation and bilingual insight.

Finally, a huge thank you goes to the English Language Program staff at Northwestern. Especially thank you to Jen Alexander and Kathleen Geraghty for their amazing peer advice throughout graduate school and THANK YOU (times infinity) to Julia Moore for her continued and unwavering support of me at all stages in this lengthy process.

DEDICATION

To Jeanne and Maria Carmela for teaching me the importance of bilingualism from a young age.

TABLE OF CONTENTS

Chapter 1. Introduction....11
1.1 Early Acquisition and Language Dominance....11
1.1.1 Early Exposure in Speech Learning....13
1.1.2 Early Exposure in Morphosyntax Learning....17
1.2 Early Exposure in Vocabulary Learning....23
1.2.1 Monolingual Vocabulary Acquisition....23
1.2.2 Bilingual Vocabulary Acquisition....27
1.3 Hypotheses and Research Questions....32
Chapter 2. Experiment 1: Speech Intelligibility....35
2.1 Methods....35
2.1.1 Participants....35
2.1.2 Talkers....35
2.1.3 Listeners....37
2.1.4 Stimuli....38
2.1.5 Procedure....39
2.2 Results....41
2.2.1 Modeling Results....41
2.2.2 English Results....42
2.2.3 Discussion of English Results....43
2.2.4 Spanish results....44
2.2.5 Discussion of Spanish Results....45
2.2.5 Acoustics Correlates of Speech Intelligibility....46
Chapter 3. Experiment 2: Lexical Decision (Reaction Time)....49
3.1 Methods....49
3.1.1 Real Word Stimuli....49
3.1.2 English Stimuli....49
3.1.3 Spanish Stimuli....51
3.1.4 Finalizing Real Word Stimuli....52
3.1.5 Non-word Stimuli....55
3.1.6 Participants....56
3.1.7 Procedure....58
3.2 Results....60
3.2.1 Modeling Results....60
3.2.2 English Results....61
3.2.3 Discussion of English Results....67
3.2.4 Spanish Results....68
3.2.5 Discussion of Spanish Results....75
Chapter 4. Experiment 3: Word Reading Time and Duration....77
4.1 Methods....77
4.1.1 Stimuli....77

4.1.2 Participants......77
4.1.3 Procedure......78
4.2 Results......79
4.2.1 Modeling Results......79
4.2.2 English Results......82
4.2.3 Discussion of English Results......87
4.2.4 Spanish Results......89
4.2.5 Discussion of Spanish Results......95
Chapter 5. Concluding Discussion......97
5.1 Summary of Results......97
5.2 Early Acquisition Versus Extended Use......101
5.3 Defining a Native Speaker......105
5.4 Processing the Non-dominant L1......108
5.5 Pedagogical Implications......111
5.6 Limitations and Future Directions......112
5.7 Conclusion......116
References......118
Appendix A: List of Experiment 1 Stimuli......129
Appendix B: List of Experiment 2 and 3 Real Word Stimuli......140
Appendix C: List of Experiment 2 and 3 Non-word Stimuli......147
Appendix D: Additional Figures from Experiment 2......151
Vita......155

LIST OF TABLES

Table 1: Number of English Items Available after Initial Selection Process 50
Table 2: Number of Spanish Items Available after Initial Selection Process................................52
Table 3: Lexical Frequency and Age of Acquisition for Items in Experiments 2 and 3 53
Table 4: Syllable Length for Items in Experiments 2 and 3 55
Table 5: Error Analysis of English Lexical Decision Results 63
Table 6: Error Analysis of Spanish Lexical Decision Results...70

LIST OF FIGURES

Figure 1: English Speech Intelligibility Scores for English controls and Heritage Speakers........43
Figure 2: Spanish Speech Intelligibility Scores for Spanish controls and Heritage Speakers.......45
Figure 3: Distribution of Stimuli in Experiments 2 and 3: Age of Acquisition and Lexical Frequency 54
Figure 4: Percent Correct Identification of English Stimuli 62
Figure 5: Age of Acquisition on English Word Reaction Time 64
Figure 6: Lexical Frequency on English Word Reaction Time 65
Figure 7: Interaction of Age of Acquisition and Group on English Word Reaction Time...........67
Figure 8: Percent Correct Identification of Spanish Stimuli..69
Figure 9: Age of Acquisition on Spanish Word Reaction Time...71
Figure 10: Lexical Frequency on Spanish Word Reaction Time..72
Figure 11: Interaction of Age of Acquisition and Group on Spanish Word Reaction Time 74
Figure 12: Interaction of Lexical Frequency and Group on Spanish Word Reaction Time 75
Figure 13: Sample Speech Segmentation for Experiment 3 80
Figure 14: Age of Acquisition and Lexical Frequency on English Word Duration 85
Figure 15: Age of Acquisition and Lexical Frequency on English Reading Time........................88
Figure 16: Age of Acquisition and Lexical Frequency on Spanish Word Duration......................92
Figure 17: Age of Acquisition and Lexical Frequency on Spanish Reading Time 95
Figure D1: Overall Reaction Time for English Controls and Heritage Speakers........................151
Figure D2: Interaction between Lexical Frequency and Group on English Word Reaction Time.... 152
Figure D3: Overall Reaction Time for Spanish Controls and Heritage Speakers 153
Figure D4: Individual Group Reaction Time Divided by Age of Acquisition and Lexical Frequency.....154

Chapter 1. Introduction

1.1 Early Acquisition and Language Dominance

Language is fascinating human ability, as infants acquire it with little to no explicit instruction from a parent (e.g., Chomsky, 1980). However, learning a second language as an adolescent or adult often results in unsuccessful areas of language acquisition such as pronunciation or morphosyntax (Reed, 1995) compared to a talker's native (often defined as first-acquired) language. Furthermore, despite this lack of explicit teaching, children are continuously exposed to language at home and in the community in which they live, ultimately resulting in their ability to acquire and produce a native language. It is still unclear which factors help or hinder this language learning process. The goal of this study is to add to the growing body of literature on language acquisition by providing insight into the way that language-learning mechanisms capitalize on the type of input children receive from an early age and that has persisted (or not) into their language use at adulthood. By exploring how some bilinguals maintain their two languages in adulthood, this study will help refine the meaning of a "native" language speaker and whether this definition of "native" must always include "first-acquired" as a necessary component.

As such, a main focus in of the present research is how language-learning mechanisms capitalize on the timing (i.e., age of acquisition) and overall amount (i.e., extended usage resulting in language dominance) of linguistic exposure. Data from monolinguals are unable to distinguish these two factors, as fluent monolinguals are necessarily defined as L1 (first language acquired)-dominant speakers. It could follow that studying bilingual speakers potentially de-confounds these two factors (early acquisition versus language dominance); however, often participants in bilingual language learning studies (e.g., Bosch, Costa, and Sebastián-Gallés, 2000; Rogers, Lister, Febo, Besing, and Abrams, 2006), while proficient in their L2 (second-acquired language), are still L1-dominant speakers. As such, it is difficult to pinpoint the source of successful (or unsuccessful) language learning in children that persists through adulthood. That is, the key element for successful language learning could be early exposure to the linguistic stimulus. Conversely, regardless of the order in which the stimulus was first acquired, extended exposure to the linguistic system resulting in language dominance could modulate language learning. As such, it is necessary to examine bilinguals who vary with respect to their L1 (first-acquired) and dominant language (i.e., bilinguals who are L2-dominant) in order to remove some of the confounding factors in language learning and better understand what types of exposure patterns language learning mechanisms capitalize on.

One such type of bilingual speaker who de-confounds these two factors is a switched-dominance bilingual for whom their dominant language is the L2 (second-acquired language), not the L1 (first-acquired language). These bilingual speakers, often referred to as heritage speakers[1], allow for a direct comparison of the effects of L1 status to those of dominant language status. While typical bilingual and monolingual populations confound these two factors (early acquisition and language dominance), switched-dominance bilinguals, or heritage speakers, dissociate them.

Switched-dominance bilingualism typically surfaces in children of immigrants (whose parents are L1-dominant in their home language) who attend school where the home language is not the primary language of instruction (Montrul, 2010a; Polinsky and Kagan, 2007) or is a minority language in the community. For example, in the United States, due to heavy immigration from Latin American countries, there are a number of Spanish-English switched-

[1] Throughout this study, L2-dominant or switched-dominance bilinguals will be referred to as heritage speakers or HS. The term "native-like" in this study will refer to a monolingual, L1-dominant speaker with no other significant influence of a second language.

dominance bilinguals whose parents are L1-dominant speakers of Spanish (e.g., Montrul, 2010a). These children begin primary schooling in the US at approximately five to six years old in which the language of instruction is English (throughout high school), thus leading to L2-dominance in English.

Such a population can help identify which areas of linguistics representation are modulated by early exposure or language dominance by examining the linguistic performance by heritage speakers compared to L1-dominant controls. Generally speaking, research on heritage speakers shows a divide between speech learning (speech production and perception) and morphosyntax learning (verbal morphology, case marking, agreement, etc.) with previous research suggesting that early exposure provides resistance against degradation to speech learning in adulthood (e.g., Au, Knightly, Jun, and Oh, 2002) while early exposure is insufficient for robust morphosyntax learning (e.g., Montrul, 2010b). However, as we will see, this divide may not be as sharp as previously believed, suggesting that both early exposure and language dominance influence the development of an overall linguistic system. The following section provides background research on current issues in heritage speaker language acquisition.

1.1.1 Early Exposure in Speech Learning

With respect to heritage speaker language learning, we see that early exposure generally provides some benefits for speech learning via resistance to degradation of the non-dominant L1 at adulthood. For example, Au, Knightly, Jun, and Oh (2002) analyzed the voiced onset times (VOTs) of productions of Spanish (non-dominant L1) plosives word-initially and Spanish morphological structures in adult Spanish-English heritage speakers (SHS; L2-dominant English). Spanish has a pre-voiced vs. short lag VOT contrast for plosives (e.g., /g/ and /k/ as in *gota—'drop' and cota—'coat (of arms)',* respectively) whereas English has a short lag vs. long lag (aspiration) VOT contrast (e.g., /p/ and /p^h/ as in ‘bail’ and ‘pail’, respectively). They found that despite being clearly dominant in their L2 (English) they were still able to produce a native-like VOT contrast (pre-voiced versus short lag) in Spanish. An L2 learner comparison group (L1-dominant English; acquired Spanish after age 14) in the study, however, was not able to produce the Spanish VOT contrast and instead produced a more English-like contrast on Spanish tokens (short lag versus long lag/aspiration). Oh, Jun, and Knightly (2003) also showed a similar effect with the Korean three-way contrastive VOT system, which contains short lag (/p/), long lag (/p^h/), and tense (/p͈/) distinction on plosives. In this study, they demonstrated that Korean-English heritage speakers (KHS; L2-dominant English) who were only passively exposed to Korean as a child (that is, their parents, who despite being L1-dominant in Korean still spoke to them in English while their parents spoke in Korean to each other), still reliably differentiated between the three-way VOT contrast in Korean (their non-dominant L1). Furthermore, the KHS levels of VOT discrimination were similar to those of monolingual Korean listeners, whereas novice learners (late English-Korean bilinguals; first acquired in high school and college; beginner level placement) could not reliably distinguish this contrast. These studies demonstrate that even when faced with limited, interrupted exposure to their L1, heritage speakers were still able to perform at L1-dominant native-like levels in some areas of speech perception and production in their L1.

Vowel spaces may also be subject to more resistance due to early exposure even in the non-dominant L1. The back vowel system of Mandarin and English contains notable distinctions with English having /u^w/ and /oʊ/ as contrastive phonemes while Mandarin distinguishes /u/ and /ɤ/ phonemically. Chang, Yao, Haynes, and Rhodes (2011) questioned whether Mandarin heritage speakers (MHS; L2-dominant English) would be able to correctly separate Mandarin and English back vowels. When compared to L1-dominant Mandarin (L2-late English) speakers and L1-dominant English (L2-late Mandarin) speakers, MHS showed greater separation of Mandarin and English vowel spaces compared to either group[2]. This result suggests that the early exposure to Mandarin was sufficient to form an independent vowel space, despite the later, dominant influence of English. Godson (2004) showed a similar effect with Western Armenian heritage speakers (WAHS; L2-dominant English) living in California. Western Armenian also contains a back vowel contrast similar to, but distinct from, the English /u/-/o/ back vowel contrast and /i/, /a/ and /ɛ/ vowel categories, which are realized in a distinct vowel space arrangement from the corresponding categories in the English vowel space. In this production study, WAHS were able to maintain the back vowel /u/ and /o/ categories as distinct from English /uʷ/ and /oʊ/ similar to native Western Armenian speakers, but failed to separate /i/, /a/ and /ɛ/ from their English /iː/, /ɑ/ and /ɛ/ equivalents. It is unclear why high and/or mid back vowels were more resistant to degradation in the heritage language in these two studies. Regardless, these studies demonstrate that some areas of heritage speaker speech learning benefitted from early (despite interrupted) exposure to language. However, they also demonstrate that there is some vulnerability in the L1 of heritage speaker speech learning as front (and the mid low /a/) vowels were not as resistant to influence from the dominant L2.

Interestingly however, in the majority of these studies (e.g., Au et al, 2002; Godson, 2004; Oh et al, 2003), stimuli were presented and measured in ideal listening and speaking conditions (e.g., in the absence of background noise). Furthermore, studies like Chang et al (2011), Godson (2004), and Oh et al (2003) focused on segment-level speech production and perception, without considering how exposure patterns affect word- and sentence-level

---

[2] These same MHS also showed greater separation of English and Mandarin plosives (voiced onset times) and fricatives (articulatory placement) compared to either bilingual group, consistent with Au et al, 2002, Oh et al, 2003.

processing. For example, it may be the case that while the KHS in Oh et al (2003) were able to correctly distinguish the three-way Korean VOT contrast in quiet listening conditions, the introduction of background noise in the signal may cause perceptual difficulties for these listeners but not for native Korean (i.e. Korean-dominant) listeners, as it has been shown that processing speech in noise can be more difficult for bilingual listeners over all (e.g., Mayo, Florentine, and Buus, 1997; Rogers et al, 2006). The extent to which early exposure protects against situations in which the linguistic system is stressed remains unclear.

To better understand the different role of early exposure in speech learning, Blasingame and Bradlow (in prep) compared the speech-in-noise recognition accuracy (speech perceptions) of Spanish heritage speakers (SHS) to late English-Spanish bilinguals and late Spanish-English bilinguals under a variety of signal enhancements and degradations in both English and Spanish. Critically, this study measured word-level accuracy (in sentence-final positions) in all three bilingual groups, rather than focusing on phoneme or segment level speech learning. They found that in favorable listening conditions (e.g., low levels of background noise), SHS had native-like levels of speech recognition accuracy (compared to the L1-dominant controls) in both languages, consistent with Au et al (2002) and Oh et al (2003) who argue that early exposure to the non-dominant language provides some benefits to speech learning. However, only the dominant language of the SHS (here, English) showed native-like resistance to degradation (as measured by sentence-final word recognition accuracy) in unfavorable listening conditions, despite Spanish (the L1) being acquired first. As such, while speech-learning mechanisms seem to capitalize on early exposure (e.g., Au et al, 2002; Rogers et al, 2006, etc.), early exposure may be insufficient to ensure resistance to adverse conditions for speech perception.

1.1.2 Early Exposure in Morphosyntax Learning

Although early exposure being beneficial to language acquisition is not a controversial statement, it is still unclear how much benefit early exposure provides to language acquisition. There is a large body of evidence, particularly on morphosyntax, that has shown early exposure to be insufficient independently for successful language acquisition (i.e., native-like performance). Au et al (2002) further demonstrated in Spanish heritage speakers (SHS; L2-dominant in English) that their morphology and gender agreement learning (productions of determiner, adjective, and noun sequences), however, patterned more similarly to L2 learners of Spanish rather than native L1 Spanish speakers. For example, Spanish requires the determiner and adjective to agree (morphologically) with the noun in both number and gender whereas English does not (e.g., 'la mesa rota'—*the broken table* (the [feminine, singular] table [feminine, singular] broken [feminine, singular] versus *'los mesa roto' (the [masculine, plural] table [feminine, singular] broken [singular, masculine]). Late English-Spanish bilinguals and SHS typically made similar morphological agreement errors (*'los mesas'), whereas L1-dominant Spanish speakers did not (Au et al, 2002). These results support the claim that some learning mechanisms (to be discussed below in this section) are more strongly influenced by language dominance, rather than early exposure for native-like function in adulthood.

Morphological agreement is not the only morpho-syntactic feature that is affected by language dominance. Montrul (2010b) showed that SHS, despite being L1 Spanish speakers, had L2 learner-like morphosyntax productions and grammaticality judgments. In Spanish, nouns in the accusative position that have [+person] as a semantic component require an obligatory case marker 'a' as in *Juan vio a Maria* ('John saw [CASE] Mary'). However, L2-dominant Spanish-English bilinguals, lacking this case marker in English, generally failed to produce this case marker and rated sentences that would be ungrammatical (**Juan vio María* ‘John saw Mary’ vs. *Juan vio a María* ‘John saw [CASE] Mary’) as acceptable Spanish sentences whereas L1-dominant Spanish speakers did not make such errors. Despite both having Spanish as the L1, the heritage speakers in this study relied on their dominant language English for grammaticality judgments rather than the L1 (Spanish). Montrul (2010b) goes on to argue that this is a case of dominant language transfer. That is, the dominant language (independent of the order of acquisition) will transfer grammatical aspects to the non-dominant language (in this study, morphological case).

Language dominance also modulates case in more morphologically rich languages. Polinsky (2008) demonstrated that Russian-English heritage speakers (RHS) failed to correctly categorize the three-way case distinction for Russian nouns compared to L1-dominant Russian speakers. In Russian, there is a complex system of case marking with four classes of nouns and six cases creating a large number of unique combinations of case on nouns (e.g., Class I-masculine-nominative: -∅), Class II-nominative: /a/; Class IV-nominative: /o/). When prompted to provide the correct case on nouns in Russian, RHS made significantly more errors compared to L1-dominant Russian speakers on gendered case agreements despite both groups acquiring Russian as their L1. Specifically, they transferred the Class I marker, which is typically phonologically null, onto nouns that required an overt phonological form of the morpheme. Polinsky (2008) argues that this is dominant language transfer from English (which has no morphological case) to Russian. However, it is unclear if this interpretation is correct, as it would be difficult, if not impossible with this data set, to measure whether RHS acquired the complex case distinction and then forgot it (i.e., language attrition) or whether they never acquired the case markers initially. An alternative interpretation is that this result could also suggest that the early exposure to Russian was insufficient to form native-like (i.e., L1-dominant) morphosyntactic representations, remaining indifferent to whether RHS acquired the complex case distinction or if they lost it in adulthood.

However, there is some conflicting evidence with respect to heritage speakers and morphological case. Korean also has a rich system of case markers on nouns indicating grammatical function in a sentence (e.g., /ga/ indicates nominative/subject of the sentence; /lul/ indicates accusative/direct object). Despite these case markers being obligatory in most cases, Chung (2018) found that L2 non-dominant Korean speakers (L1-dominant in English, which does not have morphological case) were significantly more likely to drop case markers compared to Korean heritage speakers (KHS). This result would initially suggest that the early exposure to Korean was beneficial to form case marking in KHS. However, in this study there was no comparison to L1-dominant Korean speakers such that it could be the case that the KHS pattern of case marking still failed to reach native-like (L1-dominant) levels of performance or that L1-dominant Korean speakers drop case[3] in a principled way that is distinct from KHS. However, despite this missing comparison, the fact that KHS accurately use case more than late L2 learners of Korean, suggests that the early, although limited, exposure to Korean was at least in part beneficial to morphosyntax acquisition.

Additionally, Montrul, Davidson, Fuente, and Foote (2014) found that early exposure to Spanish provided at least some benefit to morphological agreement for Spanish heritage speakers (SHS) compared to late learners of Spanish (L2 learners). In this study, L1-dominant Spanish speakers, L2-dominant SHS and late L2 learners of Spanish (L1-dominant English) were given grammatical and ungrammatical determiner-adjective-noun sequences in Spanish

[3] In some East Asian languages that contain case marking such as Japanese and Korean, case can be optionally dropped in some contexts.

(grammaticality based on morphological agreement as in *la mesa rota* where the determiner, noun and adjective all necessarily must agree on either masculine or feminine forms) and asked to perform a variety of grammaticality tasks. Despite being slower and less accurate than L1-dominant speakers of Spanish on these tasks, SHS were faster and more accurate compared to late L2-learners of Spanish on some tasks. Interestingly, in cases of non-canonical gender marking (e.g., *la calle amarilla*—'the yellow road' where *calle*—'road' does not contain the canonical –a morpheme indicating feminine), SHS were more adversely affected[4] (i.e., slower reaction time, less accurate) compared to canonical word endings (where the morpheme is consistent with grammatical gender). Critically, this effect was not as strong for late L2 learners.

Word order is also subject to dominant language influence, regardless of the order in which a talker's two languages were acquired. Albirini, Benmamoun, and Saadah (2011) showed that Arabic word order was subject to incomplete acquisition by Arabic-English switched-dominance bilinguals (AHS) due to interference from competing English structures. In some dialects of Arabic, particular sentence constructions may be Verb-Subject-Object (VSO), whereas in English, such sentences are obligatorily Subject-Verb-Object (SVO). Albirini et al (2011) found dominant language transfer from English to Arabic for AHS as the AHS did not produce the correct form of the Arabic word order. Instead, AHS produced a regularized, English-like form (SVO), even though their participants had been exposed to Arabic from childbirth whereas L1-dominant Arabic speakers correctly produced the correct VSO form. The authors interpreted this result as a failure of early exposure to protect against the dominant (L2) transfer of English. One could also question whether AHS fail to acquire this word order in

[4] The authors of this study contribute this canonical versus non-canonical difference in heritage speakers to a frequency-lag hypothesis (e.g., Gollan, Slattery, Goldenberg, van Assche, Duyck and Rayner, 2011). That is, because non-canonical forms are less frequent and HS have overall less exposure to the L1, they demonstrated poorer linguistic performance compared to L1-dominant speakers.

Arabic. However, unlike in Polinsky (2008) where Russian-English heritage speakers (RHS) produced no case marking (which is minimally consistent with an English strategy), the AHS in this study actively used the default English structure. This strategy suggests that the dominant language is still an important base for morphological processing and production. While language transfer may be at play in these studies, the current study will remain agnostic to the underlying reasons language dominance modulates certain aspects of language acquisition (i.e., whether the underlying reason is dominant language transfer or non-dominant language attrition/incomplete acquisition).

Other levels of linguistic structure also seem to rely on language dominance, independent of the order in which they were acquired. Phonotactics, the system of permissible sound sequences in a language, also seem to follow dominant rather than first language patterns. In languages such as Japanese that do not phonotactically allow consonant clusters, listeners can psycho-acoustically hallucinate the presence of a reparative vowel (or other repair strategy) even in the absence of any vowel (e.g., Dupoux, Kakehi, Hirose, Pallier, and Mehler, 1999). For example, monolingual Japanese listeners would hear an /u/ vowel between the two medial consonants in the non-word /ebzo/ perceiving it as [ebuzo], because Japanese phonotactics disallows medial consonant clusters even though no such vowel exists in the acoustic stream. These illusory vowels, as they are called, also seem to be modulated by language dominance—a conclusion that is only available when testing L2-dominant (i.e., heritage) speakers. Parlato-Oliveira, Christophe, Hirose, and Dupoux, (2010) found that language dominance modulated epenthetic repair strategies in different types of bilingual listeners. In Japanese, word final consonants are repaired via u-epenthesis, while Brazilian Portuguese repairs these consonants via i-epenthesis. For example, in Japanese, the word for ‘Christmas’ /kɹɪs.məs/ would be borrowed as /ku.ɾi.su.ma.su/; in Brazilian Portuguese, the English word ‘hot dog’ /hɑt.dɔg/ would be borrowed as /hɔ.tʃi.dɔ.gi/. As such, bilingual Japanese-Portuguese speakers must choose a strategy for phonotactic repair of either epenthetic /ɯ/ or /i/. Parlato-Oliveira et al (2010) showed that Japanese heritage speakers (L2-dominant Brazilian Portuguese) living in Brazil resorted to Brazilian Portuguese perceptual repair strategies rather than Japanese strategies. That is, when presented with a non-word cluster like /eb.zo/, they repaired it as /e.bi.zo/, a Brazilian Portuguese repair strategy. Their parents, still being L1-dominant in Japanese (and late L2 learners of Portuguese), resorted to Japanese repair strategies on the same illicit stimuli repairing it as /e.bu.zo/, a Japanese repair strategy. Despite being exposed to the same illicit consonant structures, both groups chose a different repair strategy based on their dominant, rather that first-acquired, language (recall that Japanese was the L1 for both groups).

A similar result was found in in Spanish-English bilinguals in Carlson, Goldrick, Blasingame, and Fink (2016). In Spanish, /s/ + stop consonant and nasal (/p, t, k, b, d, g, m, n/) clusters are not permitted and are repaired via prothetic /e/ at the beginning of words to resyllabify the /s/ and stop consonant or nasal. For example, a word like ‘scone’ /skoʊn/ in English would be repaired as /es.kon/ in Spanish where /es/ and /kon/ are two syllables with non-clustered onsets or offsets. The results of their study found that when presented with a non-word token like /smid/, both L1 and L2-dominant Spanish-English bilinguals displayed weakened perceptual repairs due to influence from English phonotactic structures compared to monolingual (L1-dominant) Spanish speakers who still consistently reported hearing prothetic /e/. Even more critically for this study, the L2 (English)-dominant speakers, showed even less perceptual repair of prothetic /e/ on Spanish non-words compared to Spanish-dominant bilinguals, who reported hearing prothetic /e/ relatively more. In both studies, heritage speakers (or L2-dominant speakers) did not exclusively resort to their L1's phonotactic repair strategy. Instead, heritage speakers seem to have relied on their dominant language to reconcile a non-word stimulus. These combined results suggest that language dominance may modulate phonotactic constraints more than early acquisition.

1.2 Early Exposure in Vocabulary Learning

Heritage speaker research on early exposure in language learning has largely focused on the divide between speech (production and perception) and morphosyntax (verbal morphology, case marking) learning as outlined above. However, the benefit of early exposure in vocabulary learning by heritage speakers has been less explored. One interesting feature of words is that they also contain timing features, age of acquisition (AoA) and lexical frequency, which have been shown to affect retrieval in monolinguals and bilinguals (e.g., Carroll and White, 1973). Furthermore, research in this area has shown conflicting results on whether the timing (i.e., AoA) or amount (i.e., lexical frequency) modulate successful lexical retrieval and production (e.g., Belke, Brysbaert, Meyer, and Ghyselinck, 2005 versus Zevin and Seidenberg, 2002). However, these studies are almost always conducted using either monolinguals or L1-dominant bilingual and do not allow for the unique situation of early, but limited, exposure to a language (namely, a non-dominant L1) to be studied. The following section outlines potentially conflicting roles of age of acquisition (AoA) and lexical frequency on vocabulary acquisition and retrieval and the benefit that studying switched-dominance bilinguals can provide.

1.2.1 Monolingual Vocabulary Acquisition

It has been well established in first language (L1) vocabulary acquisition studies that words learned early are recognized faster and more accurately (Carroll and White, 1973). The timing of when a word is learned affects the speed and accuracy of lexical retrieval such that early-acquired words are retrieved more quickly and accurately compared to late acquired words. Additionally, lexical frequency (here, number of occurrences of a word given a large corpus) has been shown to affect the speed and accuracy of a word's retrieval such that higher frequency words are retrieved more quickly and accurately compared to low frequency words. For example, Gerhand and Barry (1999) showed that under speeded naming conditions, both age of acquisition (AoA) and word frequency (separately) affected naming latencies. Monolingual English speakers were presented images whose lexical items varied orthogonally on age of acquisition (AoA, early versus late) and lexical frequency (high versus low) and asked to name the words for each image. They found that participants were faster at naming words that they acquired early in life and that had overall higher word frequencies. Interestingly, it should be noted that AoA had a stronger effect than word frequency on reducing naming latencies. Morrison and Ellis (2000) also demonstrated a similar effect of AoA using word estimates from real-world children's data rather than adult-rated estimates. They found that, like adult ratings of AoA, real-world childhood acquisition measurements also produced similar results. Adult native English speakers were faster in both picture naming and lexical decision tasks when the words had a relatively early AoA (based on actual childhood AoA rather than adult-reported AoA) compared to words with a late AoA. Similar to Gerhand and Barry (1999), they also found a larger effect size for AoA compared to lexical frequency in the picture-naming task, suggesting the importance of AoA as a measure of lexical retrieval.

Conversely, modeling work from Zevin and Seidenberg (2002) showed that novel word learning simulations (where no previous lexicon exists, parallel to monolingual language acquisition) rely on early exposure only in the beginning stages but that continuous exposure of individual lexical items (the amount of distribution over time) predicts successful acquisition in later stages of learning, emphasizing the importance of language dominance. Zevin and Seidenberg (2002) directly tested the two competing hypotheses of age of acquisition (AoA) and continuous exposure (CE), by manipulating the timing (early versus late) and amount (continuous versus non-continuous exposure) of exposure of artificial vocabulary items in computer-based word learning simulations (absent of any human participants). They found that in early stages of vocabulary learning, words that received exposure from the beginning (i.e, 'birth') had higher recall and recognition rates (i.e., 'better vocabulary learning') compared to words with relatively late exposures based on computer simulations. However, these differences disappeared when, at later exposures, the total amount of exposure of a given item over time was fixed between 'early' and 'late' exposed words. That is, when the amount of exposure is fixed at some endpoint, no differences between early and late exposure trajectories existed. The authors conclude that continuous exposure is a better predictor of successful vocabulary learning compared to when a word was first acquired, supporting the lexical frequency hypothesis (i.e., continuous exposure or language dominance) over the AoA hypothesis.

Ghyselinck, Lewis, and Brysbaert (2004) have shown AoA[5] to be a predictor of lexical decision processing in models that also accounted for lexical frequency[6], conflicting with the conclusions of Zevin and Seidenberg (2002). In their study, Dutch monolinguals were presented with real and non-words of Dutch and had to determine whether the stimulus was a real word in their language or not. They found that AoA was a strong predictor of reaction time such that the earlier a word was learned, the more quickly it was recognized as a real Dutch word. The authors argue that a strong version of the cumulative exposure hypothesis (i.e., age of acquisition is just

[5] For a review, see Morrison and Ellis (1995), Morrison and Ellis (2000), Morrison, Hirsh, Chappell, and Ellis (2002), and Cortese and Khanna (2007)

[6] It is important to note that Gerhand and Barry (1999) and Morrison and Ellis (2000) also show a significant effect of lexical frequency on the speed of picture naming and lexical decision (more frequent words are recognized/named more quickly)

lexical frequency in disguise) cannot account for this result and that a host of individual behavioral factors, such as subjective familiarity (i.e., how familiar a speaker is with a particular item), affect lexical access that model simulations have not taken into account thus far. Cross-linguistically, both age of acquisition and lexical frequency also affect words in other languages such as French (Bonin, Chalard, Méot, and Fayol, 2001) and Spanish (Peréz, 2007). These studies also show that age of acquisition is an independent force even when lexical frequency is controlled. However, Izura, Pérez, Agallou, Wright, Marín, Stadthagen-González, and Ellis (2011) showed that the amount of exposure to a novel word also affects its acquisition. In this study, they examined whether different types of exposure patterns (similar to Zevin and Seidenberg, 2002) affected novel vocabulary acquisition differently for English monolinguals learning a novel second language in an artificial language learning paradigm. In their study, they presented novel vocabulary items to English monolinguals and manipulated both the timing (AoA) and amount (frequency) of exposure for each word. They found that both order of acquisition and amount of exposure had measurable and independent effects on the speed and accuracy of a novel word's acquisition and retrieval.

There is also evidence that semantic competitors affect early and late words differently. For example, Belke, Brysbaert, Meyer, and Ghyselinck (2005) found that early-acquired vocabulary items are stronger competitors compared to late-acquired words. They found that when English monolinguals were given a series of semantically related words (e.g., 'duck', 'swan', 'goose', 'crane'), late-acquired vocabulary items were named more slowly in the presence of these semantic competitors compared to early-acquired ones. This result leads us to question whether the early, but not extended, exposure (the non-dominant L1) seen in heritage speakers is enough to compete with highly frequent, continuously exposed vocabulary items (the dominant L2), as well. Since then, the impact AoA (e.g., Zevin and Seidenberg, 2002) and amount of exposure (e.g., Izura et al, 2011) have on lexical access has been widely debated. More importantly, these studies establish a baseline that speakers are sensitive to features of words related to age of acquisition and lexical frequency.

1.2.2 Bilingual Vocabulary Acquisition

With respect to second language (L2) acquisition, Izura and Ellis (2002) also demonstrated that late L2 learners were sensitive to the order of acquisition (OoA) of L2-vocabulary items regardless of age of acquisition (AoA) in the L1. In their study, late Spanish-English bilinguals named early-acquired (in the classroom setting) English words more quickly and more accurately than late-acquired (in the classroom setting) words. Critically, the AoA of the semantically equivalent words in Spanish did not impact the speed or accuracy of English picture naming. This study suggests that (bilingual) vocabulary acquisition is not necessarily limited and/or solely influenced by a critical window but rather the order and amount of exposure in a given language. However, Dirix and Duyck (2017b) showed that AoA may not be fully independent in the L2 and can be affected by the L1. In this study, the eye movements of highly proficient Dutch-English bilinguals (L1 dominant in Dutch) were tracked during Dutch and English reading activities. Similar to Izura and Ellis (2002), they found that the order of acquisition (OoA) facilitated reading English words (the non-dominant L2) and replicated the AoA effect seen in L1 vocabulary acquisition studies. That is, (L2) English words with an earlier OoA were read more quickly than English words with a late OoA, with the same pattern for their L1 Dutch. However, unlike Izura and Ellis (2002), they also found that Dutch words with an early AoA also facilitated reading their translation equivalents in English. These conflicting results may be attributed to a variety of other factors well known to influence lexical processing (e.g., the bilinguals in these studies were not matched on proficiency). Setting these minor differences aside, these studies provide support for the claim that L2 learners are sensitive to age of acquisition in both languages.

It has also been well established that bilingual speakers can be slower to process vocabulary items in either language. Ivanova and Costa (2008) showed an overall slow-down in bilingual lexical retrieval. In this study, highly proficient Spanish-Catalan bilinguals (L1 dominant in Spanish) were slower to name pictures in their L2 (Catalan) compared to L1-dominant Catalan speakers, but more importantly, were slower to name pictures in their dominant L1 compared to L1-dominant, monolingual Spanish speakers. Gollan, Slattery, Goldenberg, van Assche, Duyck and Rayner (2011) also showed that bilingual speakers were more adversely affected by low frequency words compared to monolinguals. That is, even highly proficient bilingual speakers were slower to name low-frequency pictures compared to monolinguals with a variety of language pairings (namely, Spanish-English and Dutch-English). These results suggest that bilinguals, while having an overall larger vocabulary than monolinguals, may still face some challenges with respect to vocabulary processing in a particular language.

Gharibi and Boers (2017) also found that early exposure may provide some benefits to general vocabulary learning in heritage speakers, but critically, this acquisition is modulated by the age of acquisition of the second language. Specifically, the later the onset of L2 acquisition, the more successful (non-dominant) L1 vocabulary acquisition is. In this study, Gharibi and Boers (2017) compared the vocabulary sizes of (quasi-) simultaneous and sequential Farsi (Persian)-English heritage speakers (FHS, L2-dominant English) living in New Zealand in their non-dominant L1 Farsi (Persian) to L1-domnant Farsi (Persian controls). They found that the simultaneous heritage speakers (i.e., heritage speakers born in the country with the dominant L2) scored worse on Farsi vocabulary tests compared to sequential heritage speakers (i.e., heritage speakers who emigrated to the country with the dominant L2 in later childhood)[7]. Furthermore, the age at which the sequential heritage speakers immigrated to New Zealand significantly predicted vocabulary size (i.e., later arrival meant higher Farsi scores). Compared to the simultaneous speakers whose Farsi acquisition was interrupted by English even earlier, these results suggest that vocabulary learning is not as robustly protected against degradation as other areas of language learning. However, the vocabulary items tested in this study do not necessarily vary with respect to age of acquisition or lexical frequency. It could be the case that, in fact, heritage speakers are able to store certain words that have a high degree of exposure in early childhood (e.g., early-acquired or highly frequent words in the heritage language) given their limited exposure to their heritage language. However, the answer to this question is outside the scope of Gharibi and Boers (2017).

To explore whether language dominance played a role in successful vocabulary retrieval, Gollan, Montoya, Cera, and Sandoval (2008) showed that slowed down picture namings in bilinguals were exaggerated in bilinguals' non-dominant language compared to monolinguals. In this study, they found that Spanish heritage speakers (SHS; L2-dominant in English) were more adversely affected by low frequency words in Spanish (their non-dominant L1) picture naming tasks compared to monolinguals. That is, the SHS in this study were significantly slower to name low frequency words in their non-dominant language, Spanish, which in turn exaggerated the difference between high and low frequency words compared to English monolinguals[8]. This result suggests that the effect size of lexical frequency is smaller in the dominant language of

[7] Both groups received lower vocabulary scores compared to L1-dominant Farsi (Persian) speakers
[8] This study, however, lacks an L1-dominant Spanish comparison.

speakers (independent of the order in which the language was acquired), which we will see may account for some of the results in the current study. Additionally, Gollan, Ferreira, Cera and Flett (2014) demonstrated switched-dominance bilingual competition in other areas of lexical retrieval. To achieve this, Spanish tip of the tongue (TOT) states were induced in Spanish heritage speakers (SHS; L2-dominant English) when semantically and/or phonologically related English primes were present during picture naming tasks. The presence of an English prime increased TOT states in Spanish picture naming for SHS. This study suggests that the dominant language can interfere with lexical items in the non-dominant language when some similarities (either form, meaning, or structure) overlap and that vocabulary learning may also suffer despite early exposure to the language.

The effects of AoA and language dominance on vocabulary retrieval and acquisition in heritage speakers are not only limited to picture naming. In Montrul and Foote (2014), Spanish heritage speakers (L2-dominant English) and late L2 learners of Spanish (L1-dominant English; acquired Spanish after age 14) were asked to perform a speeded lexical decision task (e.g., "Is *perro* a real word of Spanish?") and to perform a translation task (i.e., "Is *gato* the translation for *cat*?") in Spanish. The stimuli varied in three ways: 1) early for SHS (i.e., early for L1 dominant Spanish speakers) and late for L2 learners (i.e., introduced late in curriculum), 2) early for both groups (i.e., early for L1 dominant Spanish speakers and early order of acquisition in curriculum for learners), and 3) late for SHS (i.e., late for L1 dominant Spanish speakers) and early for L2 learners (i.e., introduced early in the curriculum). All words were matched for lexical frequency. In both experiments, SHS were more accurate compared to L2 learners (correct identification of real words, correct rejection of non-words, and higher translation accuracy), yet L2 learners were overall faster to respond to stimuli. This slow down is unsurprising given that the stimuli were presented visually and not aurally affecting SHS reading speed in the non-dominant L1 compared to the L2 learners who had formal courses in Spanish and thus more exposure to reading Spanish (e.g., Montrul 2010b). Critically, both groups received a benefit of early-acquired words (i.e., faster response times) compared to late-acquired words, in their respective 'early' and 'late' categories (that is, early age of acquisition for SHS and early order of acquisition for L2 learners). These results, consistent with Izura and Ellis (2002) having demonstrated an AoA effect in both the L1 and L2, suggest that the early exposure to Spanish provided a small benefit to SHS lexical retrieval.

However, there were notable limitations with the stimuli in this study. In Montrul and Foote (2014), Spanish stimuli were selected from the MacArthur Bates Child Development Inventory (CDI: Fenson, Bates, Dale, Marchman, Reznick, and Thal, 2007). The CDI (and as a result the Montrul and Foote (2014) study) have some limitations that the current study seeks to eliminate. First, age of acquisition (AoA) only ranges from 0 to 36 months, which Montrul and Foote (2014) label as "early". Any of their "late" stimuli were assumed to be items not in the corpus. AoA in the CDI is determined by developmental milestones (i.e., 'all children should have these words by age X months') rather than self-reported or self-measured AoAs from a controlled study. Second, the stimuli are all matched for lexical frequency in this corpus. These limitations do not allow us to directly study late-acquired words or words with varying lexical frequencies that may influence lexical retrieval. Third, there is no direct comparison to L1 dominant Spanish speakers. The small effect size that resulted in their data may not accurately reflect how L1 dominant Spanish speakers respond to early versus late-acquired words. In other words, the diminished effect of early versus late AoA could be due to processing difficulties that SHS had or the stimuli were not strong enough to yield a robust result, yet Montrul and Foote

(2014) have no way of dissociating these two interpretations. Finally, the stimuli are presented visually rather than aurally, which resulted in L2 learner responding faster in trials compared to SHS over all. This result is unsurprising given that L2 Spanish learners are known to have higher reading proficiencies compared to SHS (e.g., Montrul 2010b). As such, more work in this area is needed to understand how certain features of words may affect heritage speakers' ability to acquire and maintain them.

To summarize, the importance of early acquisition and language dominance in heritage speakers is complex. On the one hand, they are characterized by relatively good (either better than L2 learners or similar to L1-dominant speakers) speech learning (i.e., VOT production, speech perception, vowel spaces etc.) in both the non-dominant L1 and the dominant L2. While there are some L1-processing vulnerabilities due to interrupted exposure in switched-dominance bilinguals, over all, speech-learning mechanisms seem to capitalize on early exposure to language. On the other hand, their morphosyntax and phonotactic learning show signs of L2 interference, incomplete L1 acquisition or L1 attrition. As heritage speakers (being L2-dominant) dissociate age of acquisition and amount of exposure in language acquisition, the general pattern that emerges from these studies is one in which the temporal dynamics of language learning (early onset versus extended exposure) interact in distinct ways with different levels of linguistic structure.

1.3 Hypotheses and Research Questions

The literature above has demonstrated the wide range of influence that both early and extended exposure have on language acquisition at various levels of linguistic structure and processing. Notably, in cases where early exposure is beneficial, it appears to be limited and does not cover the full range of linguistic output, whereas extended exposure to language (i.e., language dominance) appears to not be contingent on the order in which the language was acquired. However, given the previous research on speech and vocabulary learning, it appears that language-learning mechanisms for these levels of linguistic processing can capitalize on early exposure more than shown in other levels such as morphosyntax. This claim predicts better performance in the non-dominant L1 (i.e., performance on a linguistic based task that mirrors L1-dominant speakers) compared to the relatively weak performance of morphosyntax produced by heritage speakers

From this hypothesis, we also raise the following research questions:

1. Do the benefits to segmental production (i.e., pronunciation seen in Au et al, 2002, Oh et al 2003) provided by early exposure to language translate to more global aspects of speech production? While individual segments may be spared from degradation due to early exposure, does this benefit translate to more communicative functions of speech production at the sentence level? When the linguistic system is stressed (i.e., the introduction of noise), does early exposure still prevent degraded linguistic performance?
2. Having established that heritage speakers (L2-dominant) are sensitive, at least in part, to age of acquisition and lexical frequency in words (e.g., Montrul and Foote, 2014 and Gollan et al, 2011), is the early exposure to the non-dominant L1 sufficient to provide native-like (i.e., comparable to L1-dominant speakers) benefits to the retrieval of words? Additionally, is the delayed, but extended exposure resulting in language dominance in the L2 also sufficient to provide native-like benefits in vocabulary acquisition?

The results of three experiments (outlined below) will help guide the answers to these questions by indicating that early, but limited, exposure to language does not provide fully native-like benefits to language learning in difficult linguistic situations, yet extended language exposure can result in native- or near native-like performance even in situations where the linguistic system is stressed. Based on this evidence, we will conclude that early exposure to language provides only limited benefit to speech production and vocabulary learning and that extended exposure to language is necessary to provide robust native-like linguistic performance.

The study is organized in the following way. First, experiment 1 compares Spanish heritage speaker (SHS; L2-dominant English) speech intelligibility to L1-dominant Spanish and L1-dominant English speakers at two levels of noise, -4 dB and -8 dB signal-to-noise ratio (SNR), to determine if the segment-level production benefit seen in prior work translates to a sentence-level benefit in speech production. Next, experiment 2 compares lexical decision reaction times across SHS, L1-dominant Spanish, and L1-dominant English listeners. In this experiment, we explore whether heritage speakers are sensitive (i.e., reaction times affected) to the age of acquisition (AoA) and lexical frequency of words in both their non-dominant L1 and dominant L2 (Spanish and English, respectively) compared to L1-dominant listeners. Finally, experiment 3 explores whether HS are able to produce words that vary on AoA and lexical frequency in a similar manner to L1-dominant speakers. These results will then be discussed in a broader context of language acquisition and the importance (or lack thereof) of early exposure compared to language dominance.

Chapter 2. Experiment 1: Speech Intelligibility

In experiment 1, we compared the English and Spanish speech intelligibility scores of Spanish heritage speakers (L2-dominant English) to L1-dominant English and L1-dominant Spanish speakers' scores, respectively. The goal of this study was to determine whether speech production mechanisms benefit from early exposure such that speakers can reliably produce speech in their non-dominant L1 in a similar way that L1-dominant speakers produce speech in both easy and difficult talking environments. Furthermore, we question whether the delayed, but extended, exposure to the L2 was sufficient to overcome any processing difficulties incurred by later L2 acquisition.

2.1 Methods

2.1.1 Participants

The participants of interest in this study are the talkers whereas the listeners in this study simply provide speech intelligibility measurements (i.e., percent correct word identification) for each talker. However, both talker and listener groups will be described in this section.

2.1.2 Talkers

Spanish heritage speakers (SHS; n = 11) were recruited via flier on Northwestern University's campus. Their ages ranged from 18 to 22 years old. All SHS reported acquiring Spanish at birth (age 0), ensuring that Spanish was their first-acquired (from here L1 will refer to first-acquired language and L2 will refer to second-acquired language) language, and English between ages 5 and 8 years old. All 11 SHS participants were born in the US. Spanish usage at home was reported as exclusive (100 percent) during early childhood (before age 5) but falling to less than 20 percent during adulthood. Conversely, English usage during early childhood was nearly non-existent (less than 30 percent for all participants) yet between 80-100 percent usage during adulthood. Critically, SHS took no courses in high school in which the medium of instruction was any other language besides English, ensuring language dominance in English. The results section will demonstrate the switched-dominance (i.e., the L1 is no longer dominant language) nature of this particular group of bilingual talkers.

English controls (n = 11) were recruited via the Northwestern University Linguistics department subject pool. Their ages ranged from 18 to 22 years old. All English controls reported acquiring English at birth (age 0) and using English exclusively in the home during childhood. Usage of English at home and in society was reported at 100 percent during both childhood and adulthood. Critically, English controls took no courses in high school in which the medium of instruction was any other language besides English, ensuring that English is both their L1 and dominant language. English control participants with second language education (e.g., Spanish courses in high school) were included although no English control reported acquiring a second language before age 8.

Spanish controls (n = 11) were recruited via fliers on Northwestern University's campus. Participants were all graduate students at Northwestern University. Their ages ranged from 23 to 35 years old. All Spanish controls reported acquiring Spanish at birth (age 0) and using Spanish exclusively in the home during childhood. Spanish controls were born in Mexico (n = 5), Chile (n = 3), Ecuador (n = 2) and Peru (n = 1). No talker participants were from the Southern Cone (Uruguay, Paraguay, Argentina) or the Caribbean (Puerto Rico, etc.). However, Spanish controls at the time of the study used English rather extensively in their graduate school careers (ranging from 20 percent to 80 percent English usage at Northwestern University) compared to their exclusive Spanish usage during childhood (near 100 percent Spanish usage). Critically, Spanish controls took no courses in high school in which the medium of instruction was any other language besides Spanish, ensuring that Spanish is both their L1 and dominant language.

2.1.3 Listeners

English listeners (n = 44) were very similar to the English control talker group described above. Their ages ranged from 18 to 22 years old. These listeners were recruited via the Northwestern University Linguistics department subject pool. All English listeners reported acquiring English at birth (age 0) and using English exclusively in the home during childhood in the United States. Usage of English at home and in society was reported at 100 percent during both childhood and adulthood. Critically, English listeners took no courses in high school in which the medium of instruction was any other language besides English, ensuring that English is both their L1 and dominant language. English listeners with second language education (e.g., Spanish courses in high school) were included, yet no English listener reported acquiring a second language before age 8.

Similarly, Spanish listeners (n = 33) were very similar to the Spanish control talker group described above and were recruited using similar fliers around Northwestern University's campus. Their ages ranged from 23 to 35 years old. All Spanish listeners reported acquiring Spanish at birth (age 0) and using Spanish exclusively in the home during childhood. Spanish listeners were born in Chile (n = 5), Colombia (n = 6), Mexico (n = 12), Peru (n = 6), Puerto Rico (n = 1), and Venezuela (n = 2)[9]. Similarly to the Spanish talkers, Spanish listeners at the time of the study used English rather extensively in their graduate school careers (ranging from 20 percent to 80 percent English usage at Northwestern University) compared to their exclusive Spanish usage during childhood (near 100 percent Spanish usage). Critically, Spanish listeners took no courses in high school in which the medium of instruction was any other language

[9] The two participants from Puerto Rico and Paraguay were presented sentences produced by L1-dominant Spanish talkers but no heritage speakers.

besides Spanish, ensuring that Spanish is both their L1 and dominant language. No listener in either language group participated in the experiment as a talker.

2.1.4 Stimuli

The stimuli in this experiment consisted of 110 simple sentences in English and Spanish (n = 110 in each language) taken from Soli and Wong (2008) and referred to as the Hearing in Noise Test (HINT) sentences. These sentences were chosen because they were specifically adapted for audiometric testing with listeners in their respective native language when presented with additive noise. They have been normed for lexical status of words, grammatical complexity and sentence length in each respective language and have been used in a variety of cross-linguistic speech intelligibility experiments since their norming. The selection of these 110 sentences (more than 110 exist in English and Spanish) was arbitrary[10].

Stimuli were then recorded by the talker groups described above (English controls, Spanish controls, and SHS). Talkers were presented with each sentence individually and asked to produce the sentence as naturally and accurately as possible. English and Spanish[11] controls were recorded in their respective L1s while SHS were recorded in both Spanish (L1) and English (L2). English controls were given course credit for their participation. Spanish controls and SHS talkers were paid $10 an hour for their participation. The recording of these sentences typically took less than 20 minutes to complete. Talkers were instructed to read each sentence at a normal pace and to repeat any sentence that contained disfluencies (e.g., segmental and/or lexical stress errors, long pauses, word substitutions due to mis-readings). The best production of each

[10] These talkers are also a part of a larger corpus called the ALLSSTAR corpus (Bradlow, Ackerman, Burchfield, Hesterberg, Luque and Mok, 2010). As such, each participant recorded 120 HINT sentences but only the first 110 were chosen.
[11] Spanish controls produced sentences in English, as well, but these sentences were not included in the present study.

sentence was then selected, trimmed, and leveled to (a relative) 65 decibels (dB) in Praat (Boersma and Weenink, 2017).

In order to assess the speech intelligibility of each talker group and to remove any ceiling effects from listeners so as to highlight the subtle differences in speech production of these talker groups, speech shape noise was created in Praat (Boersma and Weenink, 2017) using a customized Praat script (created by the author). In this script, the average long-term average spectrum (LTAS) of each talker was measured across all sentences and then averaged across all talkers to produce a single LTAS-curve. White noise was then generated in Praat and then filtered along this LTAS curve creating speech-shaped noise customized for this experiment (i.e., noise that has follows the spectral curve of the speech stimuli).

2.1.5 Procedure

Listeners were seated in a sound attenuated booth and presented with each sentence recorded from talkers individually in their respective L1. Sentences were randomly presented to listeners over headphones using Max/MSP software (Max 7; cycling74.com). Listeners were presented with sentences blocked by either controls or SHS (e.g., English listeners heard only an entire block of SHS-produced English or English-control produced English, not mixed; the same applies for Spanish). Furthermore, the speech-shaped noise generated above was mixed with each sentence in Max/MSP at either a -4 dB signal-to-noise ratio (SNR) or a -8 dB SNR, blocked by listener. That is, listeners only heard either controls or SHS at either -4 dB SNR or -8 dB SNR (e.g., some English listener only heard SHS English sentences at -4 dB SNR while some other Spanish listener only heard Spanish-control Spanish sentences at -8 dB SNR). There was a 100 ms delay between sentences with an additional 100 ms onset of noise before the sentence was presented and a 100 ms tail of noise after the sentence ended. Each sentence produced by each talker was equally distributed amongst listeners and SNRs (e.g., English listener 1 heard 11 different English sentences from each SHS talker at one SNR) to minimize any talker-listener pairings that may result in lower speech intelligibility scores (e.g., for some arbitrary reason, talker 1 may be poorly understood by listener 8 but understood well by listener 9). This experiment equally distributed sentences across all talker-listener pairs to ensure that no talker-listener pair received more sentences than any other pair in the case that some unique combination resulted in lower intelligibility which may affect overall results. To summarize, listeners were blocked by both language group (control or SHS) and SNR (-4 dB SNR or -8 db SNR), but fully distributed across all sentence-talker-SNR combinations.

Listeners were instructed to type as many words as possible that they heard over the headphones even if it was one word. Listeners were required to type something to move on to the next sentence; as such, responses such as 'I'm not sure' or 'yo no sé' (I don't know) were allowed but then considered wholly incorrect. English listeners were given course credit for their participation and Spanish listeners received $10 an hour for their participation. Participants finished within 30 minutes. The typed output of each listener was then compared to the correct target sentence and with each word being scored as either wholly[12] correct (receiving a score of '1') or incorrect (some deviation from the correct form, receiving a score of '0'). Spanish accent marks were ignored and minor spelling errors (i.e., errors that did not result in a different word) were considered "correct". Scores were then averaged within each sentence (e.g., 5/7 words correct for sentence 103) and then averaged for each talker (total percentage of words understood across all listeners for a given talker) at each SNR to produce two speech intelligibility scores per

[12] It should be noted that morphological agreement in Spanish is much more widespread compared to English, which can result in lower scores for Spanish over all using this scoring method. For example, each word in *las niñas pequeñas*—the small girls must agree on plural (-s) and gender (-a), whereas in English this would not be the case. However, the comparison across languages is not of interest, as the stimuli are completely different.

talker (per respective language, if applicable): intelligibility at a relatively easy SNR (-4 dB SNR) and at a relatively difficult SNR (-8 dB SNR). Proportional scores were then log-odd transformed (log((p)/(1-p)) where p is the intelligibility score ranging from 0 to 1) for analysis.

2.2 Results

2.2.1 Modeling Results

The following section provides the speech intelligibility results from the above procedure (Experiment 1). All data were analyzed using the software R (R version 3.1.0, 2014) using maximal linear mixed effects regressions (LMERs) with random intercepts for participant[13]. The factors of interest, signal-to-noise ratio (SNR) and participant group (heritage speaker or control), were also contrast coded (0.5 for '-4 dB SNR' and 'control'; -0.5 for '-8 dB SNR' and 'heritage speaker', respectively) before building any models. The dependent variable (DV) was the log-odd transformed speech intelligibility score (log((p)/(1-p)) where p is the proportional intelligibility score ranging from 0 to 1. These scores were analyzed as log-odds in order to remove the non-linear relationship of incremental changes in proportions (expressed as percent words correctly identified) at various places on the scale (e.g., the difference between 52 percent and 62 percent is not equivalent to a difference between 88 percent to 98 percent). Significance was assessed using the likelihood ratio test in which the degree to which the data are fit by a full model versus a model excluding a factor or interaction of interest is measured. Final models for both English and Spanish converged with log-odd transformed speech intelligibility as the dependent variable, a two-way interaction (and both main effects) for the two factors of interest with random intercepts for participants as shown below:

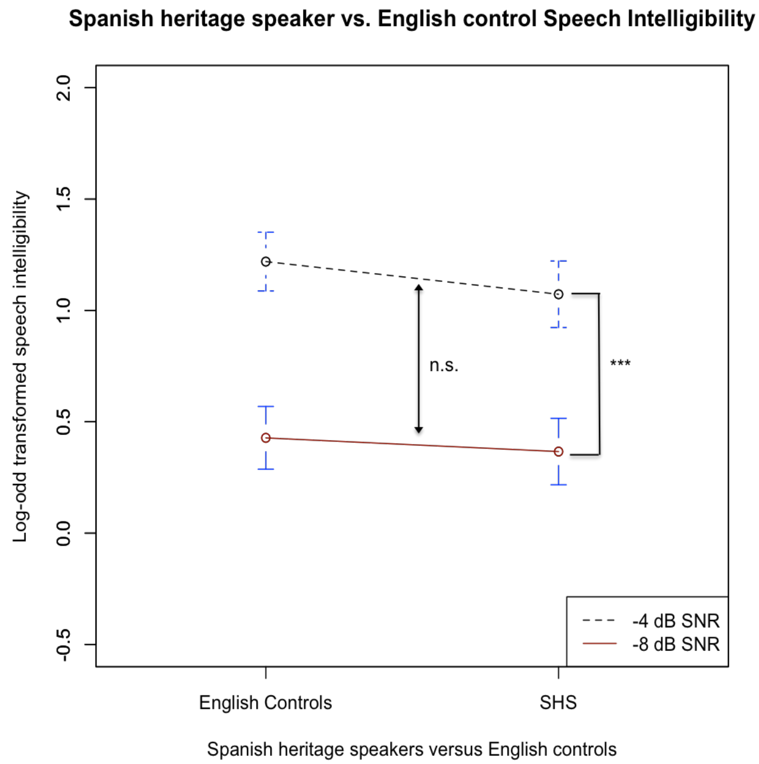
lmer(LogitScore~ Group.contrast * SNR.contrast+(1|Subject))

[13] Because there is a single intelligibility score for each talker for each SNR, a model including random slopes for SNR by talker would be over-specified and therefore, not included.

2.2.2 English Results

There was a significant main effect of signal-to-noise ratio (SNR) on speech intelligibility scores (β estimate = 0.75, standard error β = 0.03, $\chi^2(1) = 86.97$, $t = 28.79$, $p < .01$), indicating that speech intelligibility scores were significantly lower for -8 dB SNR compared to -4 dB SNR. This is a reliable effect and has been shown in many previous speech-in-noise studies (e.g., Mayo, Florentine and Buus, 1997; for a review, see general introduction). The main effect of group was not significant (β estimate = -0.1, standard error β = 0.09, $\chi^2(1) = 1.56$, $t = -1.2$, $p > .05$), indicating that the speech intelligibility scores for English controls and heritage speakers were not different. The interaction between group and SNR was also not significant (β estimate = -0.08, standard error β = 0.05, $\chi^2(1) = 2.76$, $t = -1.63$, $p = .09$), indicating that the impact of SNR on speech intelligibility scores did not differ between groups. These results are show in Figure (1).

*Figure 1: English Speech Intelligibility Scores for English controls and Heritage Speakers Interaction (not significant, indicated by the non-significant parallel lines) between signal-to-noise ratio (SNR) and group (English control versus SHS) of log-odd transformed speech intelligibility scores for English controls and Spanish heritage speakers (SHS) in English. Error bars represent 95% confidence intervals around the mean. Both SHS and English controls' speech intelligibility scores were adversely affected by SNR (significant main effect, indicated by *** on right hand side of figure) but the groups did not differ in their overall speech intelligibility scores nor was one group more adversely affected by the more difficult SNR (-8 dB SNR).*



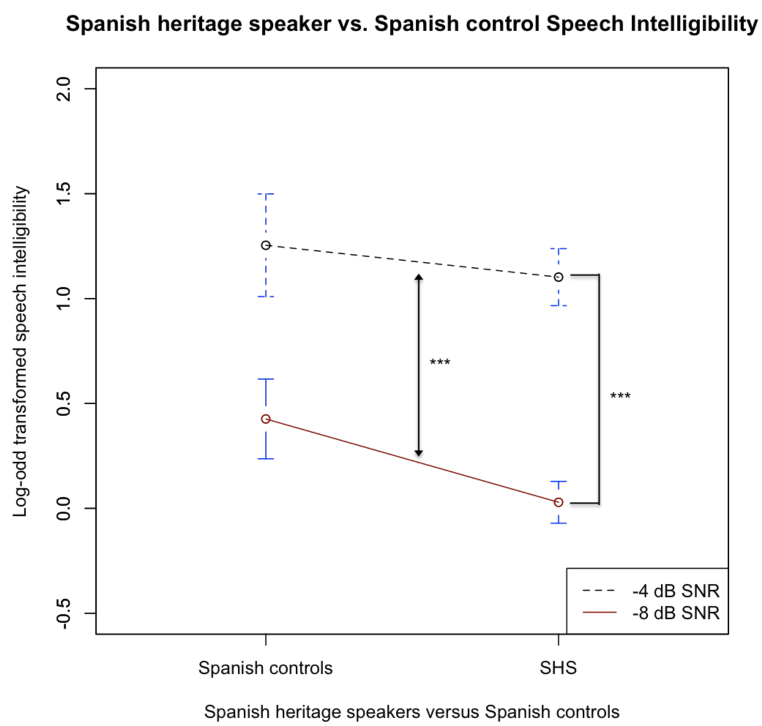
2.2.3 Discussion of English Results

The main question of this subsection was whether the delayed but extended use of language resulting in L2 (rather than L1) language dominance, was sufficient to overcome any deficits that may have been initially present in heritage speakers' delayed (but dominant) L2 speech. SHS did not show speech production vulnerabilities in their (dominant) L2 as the range of speech intelligibility scores for SHS was similar to English controls. This result is consistent with other studies that have shown L2-dominant speakers (i.e., heritage speakers) to produce native-like speech in their dominant L2 (Au, Knightly, Jun, and Oh, 2002; Montrul, 2010a; Montrul, 2010b; Oh, Jun, Knightly, and Au, 2003). The implications of these results will be discussed later in conjunction with Experiments 2 and 3.

2.2.4 Spanish Results

There was a significant main effect of signal-to-noise ratio (SNR) on speech intelligibility scores (β estimate = 0.95, standard error β = 0.05, $\chi^2(1) = 67.03$, $t = 17.47$, $p < .01$), indicating that speech intelligibility scores were significantly lower for -8 dB SNR compared to -4 dB SNR. This result is consistent with previous research (e.g., Mayo, Florentine and Buus, 1997 and many others) as well as the English results above. The main effect of group was also significant (β estimate = -0.28, standard error β = 0.1, $\chi^2(1) = 7.51$, $t = -2.87$, $p < .01$), indicating that the speech intelligibility scores for Spanish controls were significantly higher than SHS' Spanish scores. The interaction between group and SNR was also significant (β estimate = 0.25, standard error β = 0.11, $\chi^2(1) = 4.97$, $t = 2.25$, $p < .05$), indicating that SHS speech intelligibility scores were more adversely impacted by the more difficult SNR (-8 dB SNR) compared to Spanish controls. These results are show in Figure (2).

*Figure 2: Spanish Speech Intelligibility Scores for Spanish controls and Heritage Speakers Interaction (significant, indicated by the non-parallel lines ***) between signal-to-noise ratio (SNR) and group (Spanish control versus SHS) of log-odd transformed speech intelligibility scores for Spanish controls and Spanish heritage speakers (SHS) in English. Error bars represent 95% confidence intervals around the mean. Both SHS and English controls' speech intelligibility scores were adversely affected by SNR (significant main effect). Groups also differed in their overall speech intelligibility scores (Spanish controls had overall higher scores). SHS were also significantly more adversely affected by the harder SNR (-8 dB SNR).*



2.2.5 Discussion of Spanish Results

The main question of this subsection was whether early, but interrupted, acquisition was sufficient to preclude any deficits in the language acquisition process of the non-dominant L1 in adult heritage speakers, specifically with respect to speech intelligibility. In the relatively easy SNR (-4 dB SNR), SHS and Spanish controls received a similar range of intelligibility scores from native (L1-dominant) Spanish listeners. However, the significant main effect of group (Spanish controls received higher speech intelligibility scores compared to SHS) and the significant interaction between group and SNR (SHS speech intelligibility controls were more adversely affected by the more difficult -8 dB SNR compared to Spanish controls), suggests that the benefit to early exposure was insufficient to provide robst speech learning. This suggests that with regard to speech intelligibility in particular, L2-dominant bilinguals (i.e. SHS) may not be able to produce native-like speech due to interrupted L1 acquisition and replacement by a dominant L2.

It may appear that these results actually conflict with Au et al (2002) and Oh et al (2003), which have shown native-like speech learning in the non-dominant L1 of heritage speakers (for a review, see the general introduction). However, recall that in those studies noise, as a means to strain the linguistic processing system of heritage speakers, was not present. In the -4 dB SNR condition (i.e., the 'easy' listening condition), heritage speakers received comparable scores to Spanish controls in their non-dominant L1 of Spanish, consistent with Au et al (2002) and Oh et al (2003) that have demonstrated, in some cases, native-like speech learning in the heritage language (non-dominant L1). Critically, it was when the features of the non-dominant language are stressed (here, some kind of processing difficulty is introduced; later, some kind of beneficial feature of a word is absent) that differences arose between dominant L1 and non-dominant L1 speakers. The implications of this result will be further discussed in experiments 2 and 3, as well more generally at the end of these studies.

2.2.6 Acoustic Correlates of Speech Intelligibility

A natural question that may arise is what acoustic feature(s) contribute(s) to the Spanish intelligibility differences between Spanish controls and SHS in their non-dominant L1. All sentences in Experiment 1 are tightly normed for morphosyntax, meaning, and lexical status, such that the differences in intelligibility must be due to some kind of segmental (e.g., individual phonemes) or suprasegmental feature (e.g., intonation, timing, etc.). In the present study, the speech rate (number acoustic syllables divided by total time of utterance averaged across all sentences per talker), sentence duration (averaged across all sentences per talker) and average pause duration (per talker) were also measured in Praat. In Spanish, there was no significant correlation[14] between speech intelligibility and any measure of speech rate: Pearson correlation, two-tailed t-test; speech rate ($r = .075$, $t(9) = .21$, $p > .05$), sentence duration ($r = .079$, $t(9) = .22$, $p > .05$) or average pause duration ($r = -.08$, $t(9) = -.25$, $p > .05$). The results for English are similar in that none of these factors were significant correlates of speech intelligibility (speech rate ($r = -.05$, $t(9) = -.14$, $p > .05$), sentence duration ($r = .21$, $t(9) = .63$, $p > .05$), average pause duration ($r = -.008$, $t(9) = -.02$, $p > .05$)). While these global acoustic measurements did not impact speech intelligibility, further investigation is needed to determine to acoustic features that have impacted the lower speech intelligibility scores at the harder SNR in SHS' non-dominant L1.

It should also be noted that there was a high correlation between the L1 and L2 SHS speech intelligibility across both SNRs (that is, comparing the SHS English and Spanish speech intelligibility scores to one another; $r = .85$, $t(19) = 7.1$, $p < .05$), suggesting that speech intelligibility in one language may be a predictor of speech intelligibility in the other language. Furthermore, there were weak correlations of the within-SHS talker speech rate (comparing SHS English speech rate to SHS Spanish speech rate, $r = .58$, $t(9) = 2.04$, $p = .07$), sentence duration (comparing SHS English sentence duration to SHS Spanish sentence duration, $r = .62$, $t(9) = 2.4$, $p = .05$) and pause duration ($r = .61$, $t(9) = 2.2$, $p = .058$). While interesting, these results are outside the scope of the current study and require a more detailed

[14] These acoustic measurements are reported for the more difficult listening condition of -8 dB SNR. There are no correlations between speech intelligibility and any measurement at the -4 dB SNR either.

phonetic analysis. Nevertheless, note that these positive correlations of L1 and L2 temporal patterns of the SHS speech are consistent with L1-L2 speaking rate correlations reported for a much larger set of bilinguals (L1 dominant) in Bradlow, Kim and Blasingame (2017) and Bradlow, Blasingame and Lee (in prep).

One possible limitation is the SHS talkers in this study were presented with written sentences in both their dominant L2 (English) and non-dominant L1 (Spanish). As such, the global slow down observed (SHS produced slower speech rates and longer sentences in Spanish compared to English) in the non-dominant L1 (Spanish) may have lowered the SHS Spanish rates (e.g., as seen in Montrul and Foote, 2014) due to a possible greater difficulty with Spanish reading compared to English reading. The ALLSSTAR corpus (Bradlow et al, 2010) from which these stimuli were taken contains spontaneous speech recordings from SHS talkers in both languages, which may help remove confounds of reading slow-downs in heritage speakers in future work.

Chapter 3. Experiment 2: Lexical Decision (Reaction Time)

In Experiment 2, we compared the English and Spanish lexical decision reaction times (RTs) of Spanish heritage speakers (L2-dominant English) to L1-dominant English and L1-dominant Spanish speakers' RTs, respectively. The goal of this study was to determine whether lexical retrieval mechanisms benefit from early exposure such that listeners can retrieve words that vary on age of acquisition and lexical frequency in their non-dominant L1 in a similar way that L1-dominant listeners retrieve them. Furthermore, we question whether the delayed, but extended, exposure to the L2 was sufficient to overcome any processing difficulties incurred by later L2 acquisition.

3.1 Methods

3.1.1 Real Word Stimuli

The lexical items used in both this experiment (Lexical decision) and Experiment 3 (Word reading) were selected from both (separate) English and Spanish corpora. The following section outlines the criteria used to determine the final set of stimuli in both English and Spanish.

3.1.2 English Stimuli

English lexical items were drawn from the Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012) corpus. This corpus contains over 30,000 lexical items in English with oral lexical frequencies from the SUBTLEX-US (Brysbaert and New, 2009) database and Age of Acquisition (AoA) measurements obtained from L1-dominant English speaker ratings via Mechanical Turk. AoA ranged from 1 to 25 years in the English corpus. The range of ratings in this corpus is greater than the range of the Spanish corpus (see below). Lexical frequencies were back-transformed from a logarithmic scale into a base-10 numeric scale for selection. Based on the criteria laid out in Gerhand and Barry (1998), 'high' frequency words were considered above

50 occurrences per million while 'low' frequency words were considered less than 5 per million. With respect to AoA, 'early' words were acquired before the age of 5 years and 'late' words were after 8 years old.

Additionally, the following lexical items were removed from the selection process: compounds (e.g. *applesauce*), homophones and/or homonyms (e.g., *leaves* (noun) versus *leaves* (verb)), low frequency nouns whose morphologically related adjective is high frequency (e.g., target low frequency item *sadness* is morphologically high frequency adjective *sad*), diminutives, slurs and nicknames (e.g., *cutie*, *mom*, *doggie*), onomatopoeia (e.g., *bowwow*), foreignisms or borrowings (e.g., *aloha* or *mitzvah*), medical terms (e.g., *amnesia*), plural nouns (e.g., *archives*), abbreviated forms (e.g., *nuke* for *nuclear*), gerund forms (e.g., *docking*), and archaic forms (e.g., *stewardesses* for *flight attendant*). It should be noted that English words with Spanish cognates were included in the selection process. However, Spanish cognates with English, as discussed below, were not. All words used exclusively in the United Kingdom, Australia, South Africa, Canada, etc. (and not the United States) were also removed by verifying regional usage via referencing the Oxford English Dictionary.

After removing lexical items based on these criteria, the following number of words remained in each category shown in Table (1):

Table 1. Number of English Items Available after Initial Selection Process

	Early acquired	**Late acquired**
High frequency	101	34
Low frequency	62	105

3.1.3 Spanish Stimuli

The Spanish lexical items were drawn from the Alonso, Fernandez, and Díez, (2014) corpus. This corpus contains over 7,000 lexical items in Spanish with oral and printed lexical frequencies obtained from television and newspaper sources, respectively. Age of acquisition (AoA) measurements were obtained from college-level students who rated the age at which they acquired particular words ranging from ages 1 to 11. In the current studies, the oral frequencies (for Spanish) were selected as they more accurately reflected the experience heritage speakers would have with spoken, rather than written, language. Log frequencies were back-transformed to a base-10 numeric scale for selection. Again, based on the criteria laid out in Gerhand and Barry (1998), 'high' frequency words were considered above 50 occurrences per million while 'low' frequency words were considered less than 5 per million. With respect to AoA, 'early' words were acquired before the age of 5 years and 'late' words were after 8 years old.

Additionally, the following lexical items were removed from the corpus selection process: compounds (e.g., *aguanieve*—'sleet (lit. water snow)'), cognates (e.g., *aeroplano*—'airplane'), lexical items with multiple glosses (e.g., *chopo*—'poplar', *chopo*—'rifle'), lexical items without clear English glosses (e.g., place names such as *comanche*), low frequency nouns whose morphologically related adjective is high frequency (e.g., target low frequency item *desagrado*—'displeasure' related to target high frequency item *desagradable*—'unpleasant'), noun forms whose morphologically related verb is high frequency (e.g., low frequency target noun *descarga*—'(a) download)' related to high frequency verb *descargar*—'to download'), lexical items with unfamiliar glosses (e.g., *escoria*—'slag'), diminutives, slurs and/or nicknames (e.g., *cocinita*—'little kitchen' is the diminutive form of *cocina*—'kitchen'), plural nouns (e.g., *banditos*—'group of bandits'), and homophones (e.g., *asta*—'antler' and *hasta*—'until'). In the case of multiple forms of morphological gender (for a single noun), the masculine (default) form was used (e.g., *amigo*—'male friend' versus *amiga*—'female friend'). All words used exclusively in Spain (and not Latin America) were also removed by verifying regional usage via an online Spanish-English dictionary (wordreference.com).

After removing lexical items based on these criteria, the following number of words remained in each category shown in Table (2):

Table 2. Number of Spanish Items Available after Initial Selection Process

	Early acquired	**Late acquired**
High frequency	118	37
Low frequency	50	53

3.1.4 Finalizing Real Word Stimuli

The available items shown in Tables (1) and (2) were then screened further based on the range of both lexical frequency and AoA within each language. To select the final stimuli, first the median frequency and AoA along with their standard deviations were calculated within each frequency by age of acquisition cell (high-early, high-late, low-early, low-late). Then, 32 stimuli that most closely matched those numeric values were selected. In Gerhand and Barry (1998), a study that (orthogonally) examined the effects of AoA and lexical frequency on lexical decision RT (for a review, see Section I), only 8 stimuli per cell were presented to participants, thus the current study selected 16 (double) for each cell in experiment 2 and 32 (quadruple) for each cell in experiment 3. The median frequencies and standard deviations used in these selection processes are shown below in Table (3) and reflect the values of the final 128 stimuli (32 stimuli in each AoA by frequency combination) in each language (Spanish and English; 256 stimuli total). All words used in both experiments 2 and 3 were nouns (no verbs, adjectives, adverbs, etc.) and no translation equivalents across languages exist (e.g., if *dog* exists in the English stimuli, then *perro*—'dog' in Spanish was excluded and vice versa). Finally, all real words in both studies were then run through the CLEARPOND database (Marian, Chabal and Shook, 2012), which contains a variety of lexical and phonological information on Spanish and English words, to ensure lexicality (real word status), verify lexical frequency[15], and obtain orthographic and phonetic length.

Table 3. Lexical Frequency and Age of Acquisition for Items in Experiments 2 and 3
Median with standard deviation (SD) lexical frequency (log) and age of acquisition (years) for items included in Experiments 2 and 3. NB: Stimuli marked 'high frequency' are greater than 50 occurrences per million (1.699 converted to log scale) and stimuli marked 'low frequency' are less than 5 occurrences per million (.699 converted to log scale). With respect to age of acquisition, items are considered early if acquired before age 5 and late if acquired after age 8.

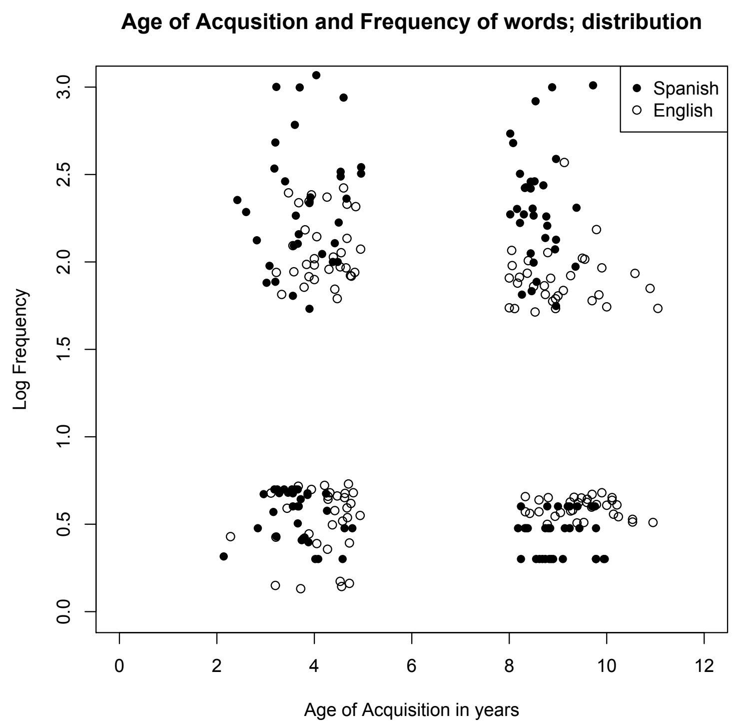
English			**Spanish**		
	Median	SD		Median	SD
AoA early	4.28	.53	*AoA early*	3.62	.67
AoA late	9.25	.76	*AoA late*	8.8	.51
Frequency high	1.94	.20	*Frequency high*	2.07	.27
Frequency low	.59	.15	*Frequency low*	.49	.14

The final set of stimuli for both languages are shown in Figure (3) below separated by log frequency and age of acquisition to ensure that stimuli groups do not overlap in numeric values. For all real word and non-word stimuli in both English and Spanish with their corresponding age

[15] Lexical frequency was only checked against the Spanish stimuli in Alonso et al (2014) as the lexical frequencies from CLEARPOND are the same as the frequency values in the Kuperman et al (2012) corpus. In cases of frequency differences (e.g., CLEARPOND has 9 per million and Alonso et al (2014) has 5 per million, stimuli were replaced until both corpora agreed).

of acquisition, lexical frequency, number of syllables, phonological length, and orthographic length (all where applicable), please see the appendix (section 2).

Figure 3. Distribution of Stimuli in Experiments 2 and 3.
The x-axis represents age of acquisition in years while the y-axis represents log frequency of lexical items. These stimuli were selected to ensure no overlap across age of acquisition (AoA) and lexical frequency



Syllable length was also measured using both the English and Spanish corpora. Although controlled for within each language, syllable length was not controlled across languages due to different phonological structures in English and Spanish, which could ultimately affect syllable length. For example, English allows complex onsets with up to three phonemes such as /str/ (e.g., *street*) and /spl/ (e.g., *splash*) while Spanish only allows (limited) complex onsets with up to two phonemes such as /fl/ (e.g., *flor*—'flower') or /gr/ (e.g., *grande*—'big'). Furthermore, Spanish codas are considerably more restricted than English codas only allowing the single consonants

/n/ (e.g., *alacrán*—'scorpion'), /l/ (e.g., *pañal*—'diaper'), /r/ (e.g., *hablar*—'to speak'), /d/ (e.g., *verdad*—'truth'), and the plural /s/ (e.g., *gatos*—'cats'). These reduced consonant clusters in Spanish onsets and codas significantly affected syllable length (two-tailed, unpaired t-test $t(254) = 8.25$, $p < .01$), indicating that Spanish words contained on average more syllables than English words. There were no significant differences in syllable length (all comparisons Bonferroni-corrected; two-tailed, unpaired t-tests $t(30) < 2$, $p > .05$) within each language across different stimuli types (early-high, early-low, late-high, late-low). These values are shown in Table (4).

Table 4. Median Syllable Length and Standard Deviation (SD) of English and Spanish Stimuli

English			**Spanish**		
	Syllables	SD		Syllables	SD
AoA early	1.63	.70	*AoA early*	2.61	.68
AoA late	2.29	.94	*AoA late*	2.95	.81
Frequency high	1.81	.87	*Frequency high*	2.63	.63
Frequency low	2.11	.89	*Frequency low*	2.94	.85

3.1.5 Non-word Stimuli

In experiment 2 (lexical decision reaction time), 64 real words of English and/or Spanish were presented orally to participants. Additionally, 64 non-word fillers of English and/or Spanish were also presented to participants (see the Procedure subsection below for a detailed description of the presentation process). This section explains the creation and selection process of non-word filler stimuli.

Non-word fillers in experiment 2 were created using "Wuggy", a non-word generator (Keuleers and Brysbaert, 2010). This stimuli generator can create Spanish and English (and other languages) non-words using a target input of real words in the respective language. The output creates stimuli that can be (user optional) set to match length, sub-syllabic length, syllabic segments, and transitional frequencies. For example, if the input is ‘grape’, a possible output is ‘crace’ (where the segments /gɹeɪp/ and /kɹeɪs/ have similar lengths, syllabic segments and transitional probabilities). 64 real words of English and Spanish were input into Wuggy (words not present in experiment 2 to avoid high lexical competition). While Wuggy allows users to select as many non-words per real word as desired (e.g., 1-100 non-words per real word), the current study selected three output non-words per every input real word. Words were matched for length of sub-syllabic segments, letter length, and transitional frequencies. The resulting output was 192 non-words, which were narrowed to 64 in each language. All homophones with real words, non-occurring spellings (‘vv’ as in ‘nuvvet’) and odd letter combinations ‘tl’, etc.) were removed. Non-words were then entered into CLEARPOND database (Marian, Chabal and Shook, 2012) to check the lexical status (to ensure it was not a real word of either English or Spanish) and number of letters (orthographic length).

3.1.6 Participants

The same participants took part in both experiments 2 and 3. This section describes their Spanish and/or English acquisition process.

Spanish heritage speakers (SHS: n = 20) were recruited via fliers on Northwestern University’s campus. Their ages ranged from 18 to 29 years old[16]. All SHS reported acquiring Spanish at birth (age 0), ensuring that Spanish was their first-acquired (from here L1 will refer to first-acquired language and L2 will refer to second-acquired language) language, and English

[16] Two SHS participants did not provide an age. Two SHS participants’ data were replaced with two new SHS due to their relative slow reading times in Spanish (time to completion in Spanish was more than double any other participant).

between ages 5 and 7[17] years old. Of the 20 SHS participants, 17 were born in the US and 3 were born in Mexico. Spanish usage at home was reported as exclusive (100 percent) during early childhood (before age 5) but falling to less than 20 percent during adulthood. Conversely, English usage during early childhood was nearly non-existent (less than 30 percent for all participants) yet between 80-100 percent usage during adulthood. Critically, SHS took no courses in high school in which the medium of instruction was any other language besides English, ensuring language dominance in English. The results section will demonstrate the switched-dominance (i.e., the L1 is no longer dominant language) nature of this particular group of bilingual speakers.

English controls (n = 20) were recruited via the Northwestern University Linguistics department subject pool. Their ages ranged from 18 to 22 years old. All English controls reported acquiring English at birth (age 0) and using English exclusively in the home during childhood. Usage of English at home and in society was reported at 100 percent during both childhood and adulthood. Critically, English controls took no courses in high school in which the medium of instruction was any other language besides English, ensuring that English is both their L1 and dominant language. As no English controls were presented with Spanish stimuli and no Spanish stimuli were translation equivalents of English stimuli, English control participants with second language education (e.g., Spanish courses in high school) were included. No English controls reported acquiring a second language before age 8.

Spanish controls (n = 20) were recruited via fliers on Northwestern University's campus. Participants were mainly graduate students, with some family members (spouses, siblings, etc.) also participating in the study. Their ages ranged from 23 to 36 years old. All Spanish controls

[17] One SHS participant reported learning English at age 3

reported acquiring Spanish at birth (age 0) and using Spanish exclusively in the home during childhood. Spanish controls were born primarily in Mexico (n = 16) and Chile (n = 3). One participant was born in Argentina. However, Spanish controls at the time of the study used English rather extensively in their graduate school careers (ranging from 20 percent to 80 percent English usage at Northwestern University) compared to their exclusive Spanish usage during childhood (near 100 percent Spanish usage). Critically, Spanish controls took no courses in high school in which the medium of instruction was any other language besides Spanish, ensuring that Spanish is both their L1 and dominant language.

3.1.7 Procedure

This section describes the procedures for participants in experiment 2.

A limitation of a similar study in which the lexical decision RT of SHS were measured (Montrul and Foote, 2014) is that stimuli were presented visually (written form) to both L2 learners of Spanish and Spanish heritage speakers (SHS). This visual presentation can (and did in fact in their study) cause a global slow-down in SHS due to their unfamiliarity with Spanish heritage reading and writing (Montrul, 2010a; Montrul and Foote, 2014). As such, all stimuli were presented aurally to participants (e.g., Taft, 1986) to avoid confounds with reading slow downs.

Stimuli were recorded in a sound attenuated booth using a Shure SM81 Condenser Handheld microphone. The English stimuli were recorded by an L1-dominant English speaker (female, age 26) who reported acquiring English at birth with no second language exposure until high school. The Spanish stimuli were recorded by an L1-dominant Spanish speaker (female, age 28) from Mexico D.F., Mexico who reported acquiring Spanish at birth with no second language exposure until high school. Each talker was instructed to read each word out loud at least three times. In the cases of non-word pronunciations[18], the Wuggy output was explained to the talkers beforehand. Both talkers had knowledge of linguistic pronunciation symbols such as the International Phonetic Alphabet (IPA). The best of each trial of recorded stimuli were then selected, trimmed, and leveled to (a relative) 65 dB in Praat (Boersma and Weenink, 2017). Of the recorded stimuli, one half of the 128 stimuli in each language were presented in experiment 2 to participants (n = 64 in each language) while all 128 stimuli were presented in experiment 3 to participants. These 64 lexical items were chosen due to their homogenous nature within each cell type (words closest to the median AoA, lexical frequency, etc.) and consistency across corpora (e.g., words with equivalent lexical frequencies in both the Alonso et al (2014) corpus and Marian et al (2012) corpus were included in experiment 2). While all (n = 64) non-word fillers appeared in experiment 2, they were not present in experiment 3.

All 128 stimuli (n = 64 real words and non-words each) were then presented to participants over headphones using Superlab (Version 5, Cedrus Corporation; San Pedro, CA). Participants were instructed that they would hear a series of audio recordings in either English or Spanish (SHS completed the task in both languages) and had to decide whether the recording they heard was a real word of the respective language. Participants were instructed to respond as quickly and accurately as possible and were given examples of real and non-real words of English and/or Spanish (e.g., ‘mug’ (real) and ‘murg’ (non-word) in English and *pierna*—“leg” (real) and *pernala*—(non-word) in Spanish). Participants could press either a labeled “yes” button (or “sí”) or “no” button to make their decision during the task. An RB-730 Response Pad (Cedrus Corporation; San Pedro, CA) button box recorded reaction time (RT) in milliseconds (ms) with button 1 being labeled “yes” or “sí” and button 7 being labeled “no”. Words were

[18] The English talker had to re-record several stimuli due to less transparent English orthography (e.g., ‘rint’ should be pronounced like /ɹɪnt/ rather than /ɹɑɪnt/) compared to the Spanish talker who had less difficulty due to the more transparent nature of Spanish orthography.

randomized for each participant independently and presented one at a time with a cross-hair fixation point on the screen during each trial with a 250 ms delay between trials. English controls were only presented with English real words and non-words. Spanish controls were only presented with Spanish real words and non-words. Spanish heritage speakers (SHS) were first presented with either English or Spanish and then presented with the other language (either English or Spanish, counterbalanced across SHS such that half of SHS (n = 10) were exposed to Spanish stimuli first and the other half were exposed English stimuli first). Participants finished the lexical decision task in less than 10 minutes in their respective languages. SHS completed both the lexical decision task and reading task (described in Experiment 3) in one language and then were given a short 10-15 minute break before completing both tasks in the other language. English controls were given course credit for their participation. Spanish controls and SHS were paid $10 per hour for their participation. The output of the reaction time labeled the participants' responses (either yes or no), the stimulus, age of acquisition (either 'early' or 'late'), lexical frequency (either 'high' or 'low'), reaction time (RT: beginning of stimulus subtracted from button press) in milliseconds, and word type (whether word is real or filler).

3.2 Results

3.2.1 Modeling Results

The following section provides the reaction time (RT) results from the above procedure (Experiment 2). All data were analyzed using the software R (R version 3.1.0, 2014) using maximal linear mixed effects regressions (LMERs) with random intercepts for participant and lexical items as well as random slopes for the three factors of interest: age of acquisition (AoA: early or late), lexical frequency (high or low) and participant group (heritage speaker or control) by subject (e.g., Barr, Levy, Scheepers and Tily, 2012). These three factors (AoA, lexical frequency and participant group) were also contrast coded (0.5 for 'high', 'early', and 'control'; -0.5 for 'low', 'late' and 'heritage speaker', respectively) before building any models. The dependent variable RT was transformed to a log scale and then any data more than 2.5 standard deviations[19] above the mean were excluded[20] from the final analyses. Additionally, only correctly identified words were analyzed (no fillers, no false negatives; In English, percent correct real-word identification, controls: 1264 out of 1280 trials (98.8%) and SHS: 1261 out of 1280 trials (98.5%). In Spanish, percent correct real-word identification, controls: 1232 out of 1280 trials (96.3%) and SHS: 1032 out of 1280 trials (80.6%)). Significance was assessed using the likelihood ratio test in which the degree to which the data are fit by a full model versus a model excluding a factor or interaction of interest is measured. Final models for both English and Spanish converged with RT as the dependent variable, a three-way interaction (and all subsequent two-way interactions and main effects) for the three factors of interest with random intercepts for participants and lexical items with random slopes for the interaction between AoA and frequency by subject was as shown below:

lmer(LogRT~ Group.contrasts* Frequency.contrasts* AoA.contrasts+(1+Frequency.contrasts*AoA.contrasts|Subject)+(1|Stimulus))

3.2.2 English Results

Main Effects:

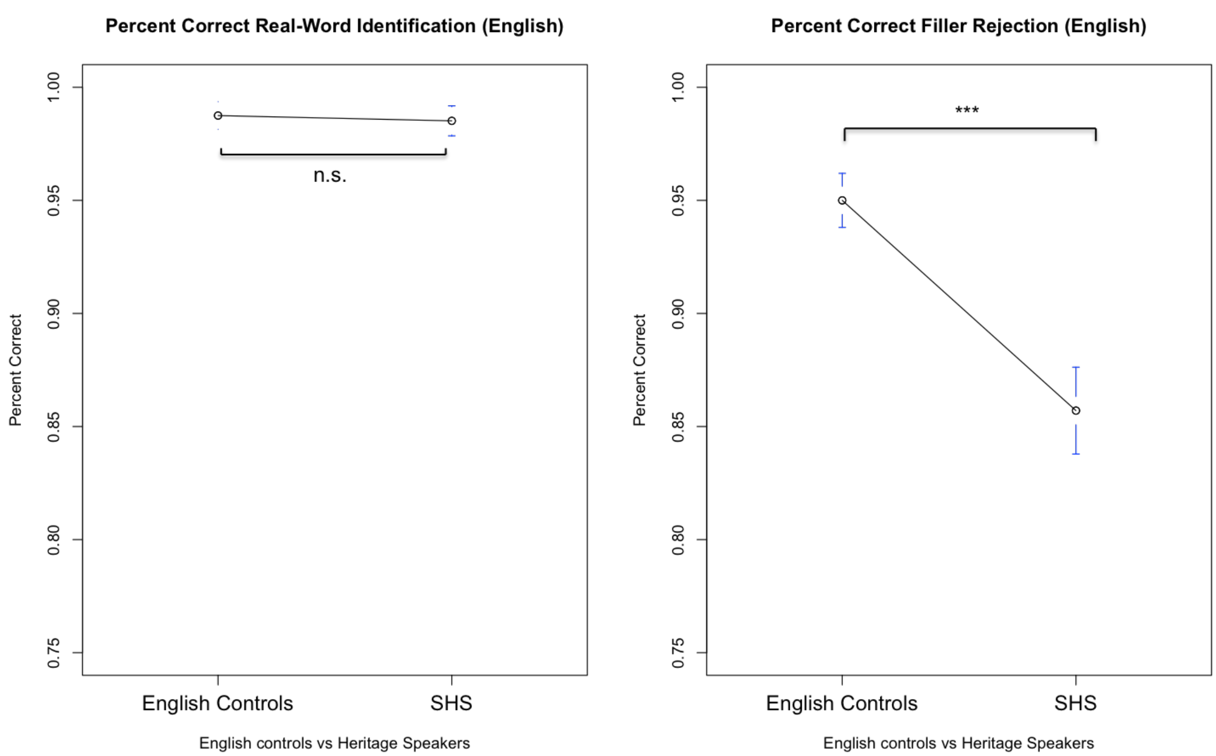
1. Group

With respect to accuracy, both groups were highly accurate at identifying real-word items (English controls: 98.75%, SHS: 98.5%; log-odds transformed overall scores; β estimate = 0.003, standard error β = 0.004, $t(38) = 0.88$, $p > .05$). Interestingly, English controls had more accurate

[19] As RT necessarily has a right-tail, only values above this threshold were excluded
[20] Models were run with the original set of data kept in and did not significantly alter results.

identification (correct rejection) of filler items in English (95%) compared to SHS controls (85%). This difference was significant (log-odd transformed overall scores; β estimate = 0.09, standard error β = 0.03, $t(38) = 2.87$, $p < .05$), indicating that English controls more accurately identified non-words of English. These results are shown in Figure (4) below. There was no main effect[21] of group (β estimate = 0.05, standard error β = 0.03, $\chi^2(1) = 1.59$, $t = 1.4$, $p > .05$) on RT suggesting that the RT to real-word stimuli was similar for English controls (median = 6.77 (sd = .20)) and SHS (median = 6.73 (sd = .28)). Table (5) provides descriptive statistics of the errors from each participant group (English controls versus SHS) by AoA and lexical frequency.

Figure 4. Percent Correct Identification of English Stimuli
Percent correct word identification (left) for English controls and SHS—no significant differences between groups. However, percent correct filler (right) rejection (correctly responding 'no' when presented with a filler) shows that English controls are significantly more accurate compared to SHS in correctly identifying non-words of English. Error bars represent 95% confidence intervals around the mean.



[21] All non significant figures for this experiment can be found in the appendix (section 4).

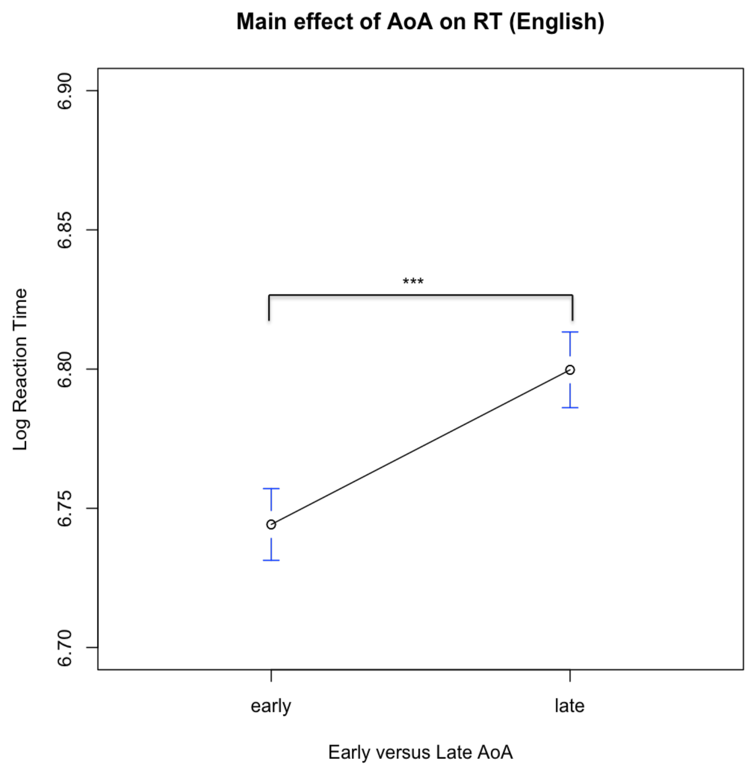
Table 5. Error Analysis of English Lexical Decision Results
Each cell contains the total number of errors out of 1,280 total trials for each age of acquisition (AoA) and lexical frequency (LF) combination (early AoA-high frequency, late AoA-high frequency, early AoA-low frequency, late AoA-low frequency) across all participants

	English controls		**Spanish heritage speakers (English)**	
	High LF	*Low LF*	*High LF*	*Low LF*
Early AoA	1 (.07%)	5 (.39%)	3 (.23%)	7 (.55%)
Late AoA	3 (.23%)	7 (.55%)	2 (.16%)	7 (.55%)

2. Age of Acquisition (AoA)

The data in Figure (5) show the effect of age of acquisition (AoA) on reaction time (RT). There was significant main effect of AoA on RT for English stimuli across groups (β estimate = -0.05, standard error β = 0.02, $\chi^2(1) = 6.29$, $t = -2.5$, $p < .05$) indicating that early-acquired words are recognized more quickly than late-acquired words. This result replicates previous findings (Carroll and White, 1973; Morrison and Ellis, 1995; Morrison and Ellis, 2000; Ghyselinck, Lewis, and Brysbaert, 2004) that have shown faster decision times with early-acquired words.

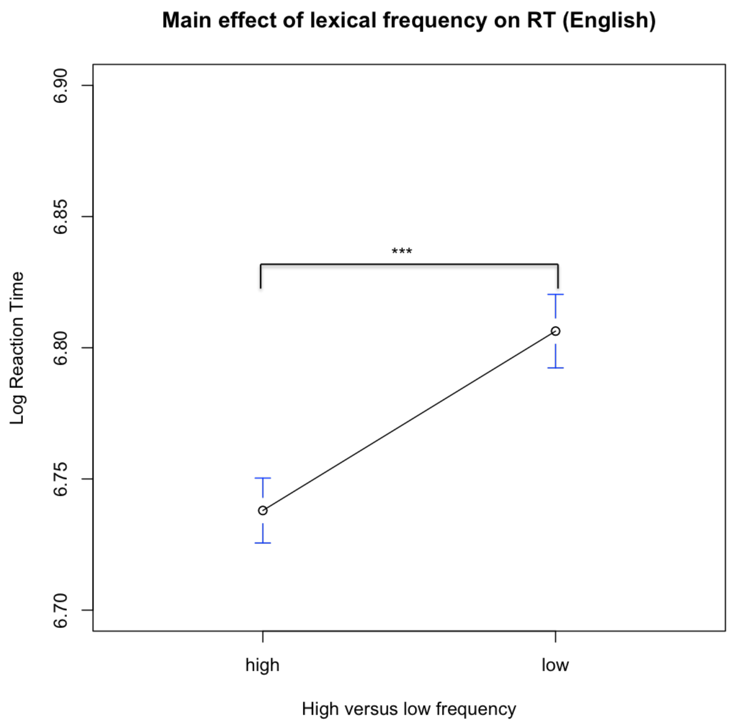
Figure 5. Age of Acquisition on English Word Reaction Time
Main effect of age of acquisition (AoA) on log reaction time (RT) for English controls and Spanish heritage speakers (SHS) in English. Error bars represent 95% confidence intervals around the mean. Participants responded more quickly to early-acquired words compared to late-acquired words.



3. Lexical Frequency

The data in Figure (6) show the effect of lexical frequency on reaction time (RT). There was a significant main effect of lexical frequency on RT (β estimate = -0.06, standard error β = 0.02, $\chi^2(1) = 7.15$, t = -2.7, p < .01), indicating that participants responded to high frequency words more quickly than low frequency words. These results also replicate previous findings (Lewis, Gerhand, and Ellis, 2000; Morrison and Ellis, 2000), which have shown that frequency is an independent force (from AoA) on lexical decision RT.

Figure 6. Lexical Frequency on English Word Reaction Time
Main effect of lexical frequency on log reaction time (RT) for English controls and Spanish heritage speakers (SHS) in English. Error bars represent 95% confidence intervals around the mean. Participants responded significantly more quickly to high frequency words compared to low frequency words.



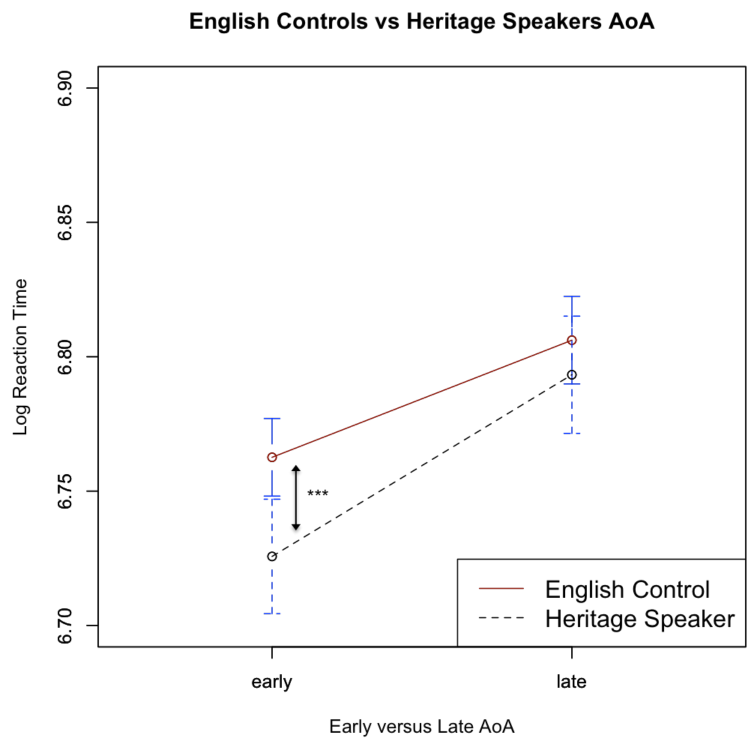
Interactions:

This section will now analyze the interactions between AoA and group (English control versus SHS) and lexical frequency and group. Critically, English controls (L1-dominant) acquired English before their SHS peers (L2-dominant). However, English controls and SHS are both English-dominant (regardless of L1 or L2 status), due to extended usage (and exposure from the community) throughout adulthood. We predict that this extended usage will benefit lexical retrieval, predicting no differences between English controls and SHS in this task.

The three-way interaction between AoA, lexical frequency, and group was not significant (β estimate = -0.04, standard error β = 0.02, $\chi^2(1) = 2.31$, $t = -1.5$, $p > .05$). Additionally, the interaction between AoA and lexical frequency was not significant (β estimate = -0.01, standard error β = 0.04, $\chi^2(1) = .11$, t = -.3, p > .05). There have been multiple conflicting results on whether AoA and lexical frequency interact (e.g., simulations from Zevin and Seidenberg (2002) suggest that the benefit of early AoA on words diminishes as lexical frequency increases compared to Lewis, Gerhand, and Ellis (2000) who claim that AoA and lexical frequency have independent forces on word retrieval). These results suggest that these forces are indeed independent on lexical retrieval.

There was a small, but significant interaction between AoA and group such that, surprisingly, SHS had faster RT to early-acquired words compared to English controls (β estimate = 0.03, standard error β = 0.01, $\chi^2(1) = 5.63$, t = 2.4, p < .05). This result is actually not predicted, as SHS are not predicted to respond faster to early-acquired words in the SHS L2 compared to English controls (English = L1). This result is shown in Figure (7).

*Figure 7. Interaction between age of acquisition (AoA) and group (English control versus SHS) on log reaction time (RT) for English controls and Spanish heritage speakers (SHS) in English. Error bars represent 95% confidence intervals around the mean. SHS participants responded significantly more quickly to early-acquired words compared to English controls (indicated by the non-parallel lines ***). However, English controls and SHS responded equally to late-acquired words.*



The interaction between lexical frequency (high versus low) and group (English control versus SHS) was not significant (β estimate = -0.001, standard error β = 0.01, $\chi^2(1) = 0.01$, t = -0.01, p > .05). This result is consistent with the prediction that despite a delay in initial acquisition to English, the extended use of English was sufficient to overcome any deficits by delayed acquisition.

3.2.3 Discussion of English Results

The main question of this subsection is whether the delayed, but frequent, use of English is sufficient to overcome any deficits that may have been initially present in heritage speakers' language acquisition process. Specifically, we want to know whether heritage speakers who are dominant, but with delayed onset, in English, are able to recognize real words of English as quickly as their early-onset and dominant English control peers. The main effects of AoA and lexical frequency are consistent with previous studies (e.g., Lewis, Gerhand, and Ellis, 2000; Morrison and Ellis, 2000) showing that L1-dominant speakers are sensitive to these lexical features. More interestingly, we have shown that any deficits that may arise in lexical retrieval due to a delay in linguistic exposure can be erased with enough frequent exposure. That is, SHS pattern like L1 English controls with respect to lexical frequency showing no significant differences in RT between groups on high and low frequency words. This result is consistent with studies such as Montrul (2010b) that argue that extended exposure to language is necessary for native-like performance. Curiously, SHS were slightly faster at responding to early-acquired words, which is not easily explained by any hypothesis in this study. The implications of these results will be discussed further below.

3.2.4 Spanish Results

Main Effects:

1. Group

First, individual group (Spanish controls versus Spanish heritage speakers, SHS) data were analyzed. These data are shown in figure (6). Interestingly, the main effect of group on RT was not significant (β estimate = -0.06, standard error β = 0.03, $\chi^2(1) = 2.63$, t = -2.3, p > .05). However, it should be noted that SHS are significantly less accurate compared to Spanish controls in both identification of real words[22] (SHS: 80.6% correct real-word identification; Spanish control: 96.25% correct real-word identification; log-odd transformed overall scores; β

[22] Recall that only correctly-identified real-word trials are included in RT analyses

estimate = -0.16, standard error β = 0.02, t(38) = -7.99, p < .05) and filler items (SHS: 86% correct filler-identification/real-word rejection; Spanish control: 93% correct filler-identification/real-word rejection; log-odd transformed overall scores; β estimate = -0.07, standard error β = 0.03, t(38) = 2.56, p < .05). This difference in lexical decision accuracy is consistent with the designation of this group of Spanish heritage speakers as switched-dominance bilinguals (i.e. for whom Spanish is the L1 (first acquired) but non-dominant language). These accuracy results can be seen in Figure (8) below. Table (6) provides descriptive statistics of the errors from each participant group (Spanish controls versus SHS) by AoA and lexical frequency.

Figure 8. Percent Correct Identification of Spanish Stimuli
Percent correct word identification (left) for Spanish controls and SHS—Spanish controls are significantly more accurate at correctly identifying Spanish words. Percent correct filler (right) rejection (correctly responding 'no' when presented with a filler) shows that Spanish controls are also significantly more accurate compared to SHS in identifying non-words of Spanish. Error bars represent 95% confidence intervals around the mean.

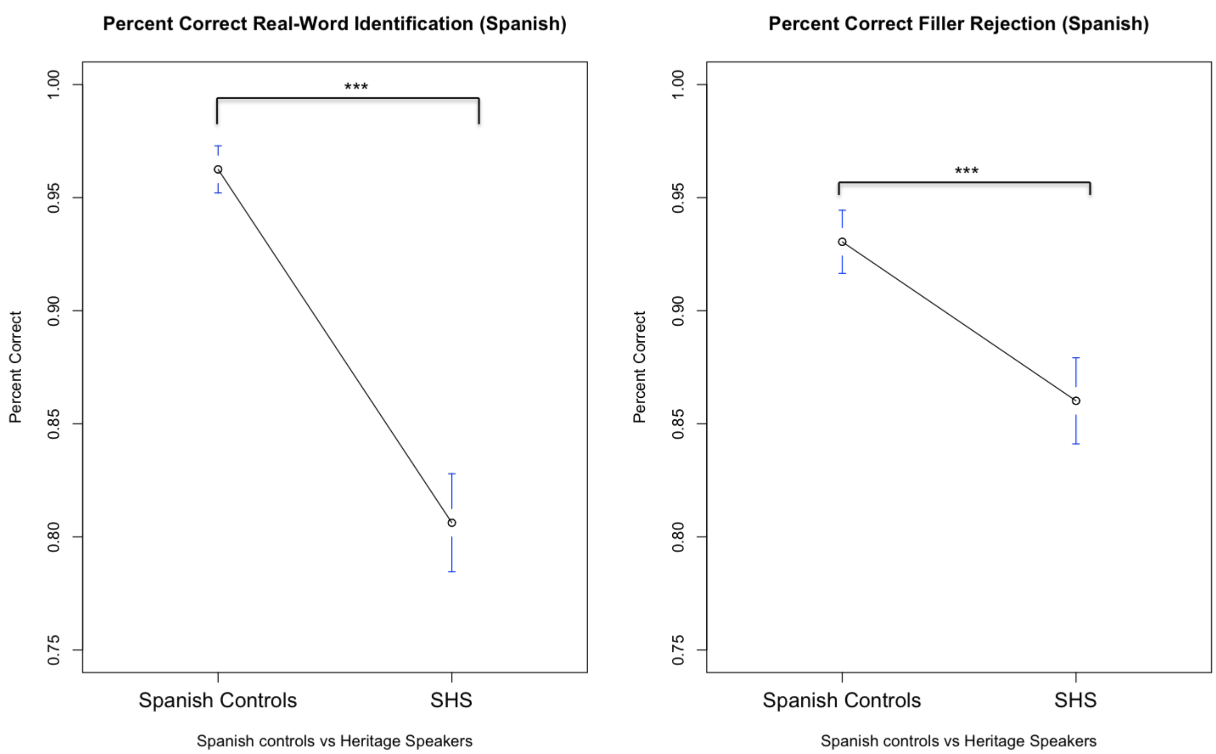


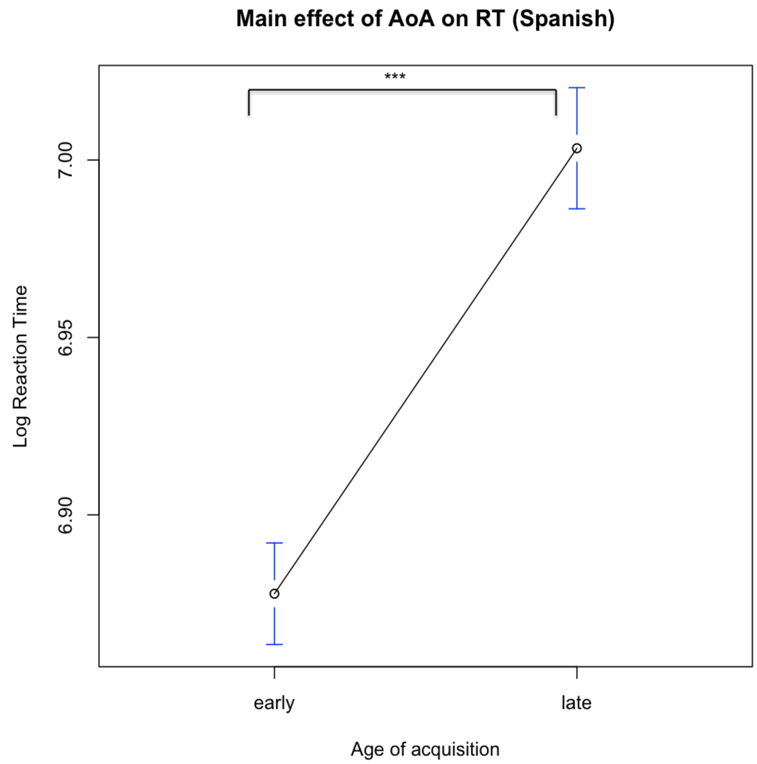
Table 6. Error Analysis of Spanish Lexical Decision Results
Each cell contains the total number of errors out of 1,280 total trials for each age of acquisition (AoA) and lexical frequency (LF) combination (early AoA-high frequency, late AoA-high frequency, early AoA-low frequency, late AoA-low frequency) across all participants

	Spanish controls		**Spanish heritage speakers (Spanish)**	
	High LF	*Low LF*	*High LF*	*Low LF*
Early AoA	3 (.23%)	5 (.39%)	2 (.16%)	60 (4.7%)
Late AoA	12 (.93%)	28 (2.19%)	23 (1.8%)	163 (12.73%)

2. Age of Acquisition (AoA)

The data in Figure (9) show the effect of age of acquisition (AoA) on reaction time (RT). There was a significant main effect of AoA on RT for Spanish stimuli across groups (β estimate = -0.12, standard error $\beta = 0.02$, $\chi^2(1) = 22.6$, $t = -5.2$, $p < .01$) indicating that early-acquired words are recognized more quickly than late-acquired words. This result replicates previous findings (Carroll and White, 1973; Ghyselinck, Lewis, and Brysbaert, 2004; Morrison and Ellis, 1995; Morrison and Ellis, 2000, as above).

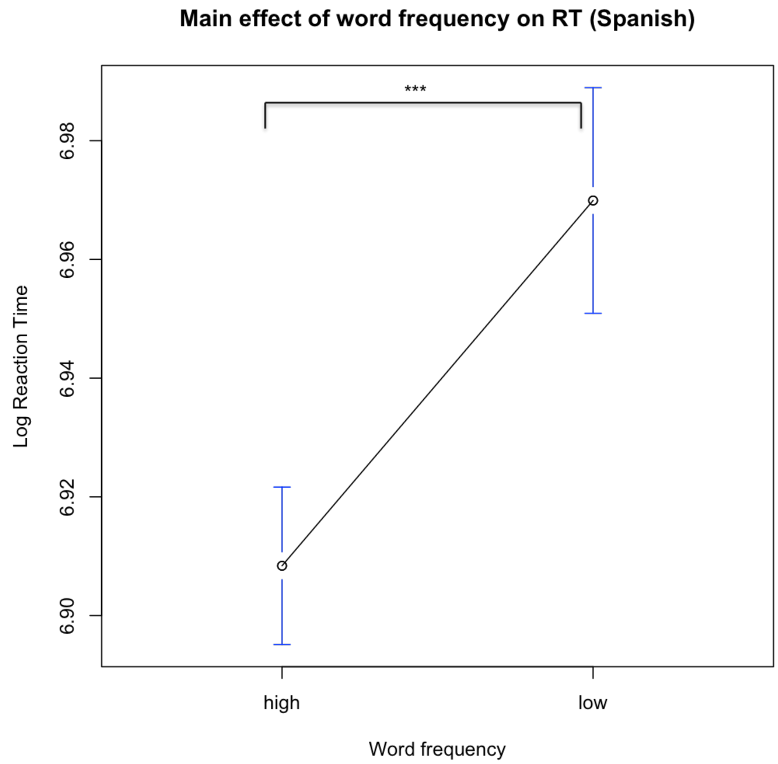
Figure 9. Age of Acquisition on Spanish Word Reaction Time
Main effect of age of acquisition (AoA) on log reaction time (RT) for Spanish controls and Spanish heritage speakers (SHS) in Spanish. Error bars represent 95% confidence intervals around the mean. Participants responded significantly more quickly to early-acquired words compared to late-acquired words.



3. Lexical Frequency

The data in Figure (10) show the effect of lexical frequency on reaction time (RT). There was a significant main effect of lexical frequency on RT (β estimate = -0.06, standard error β = 0.03, $\chi^2(1) = 4.9$, $t = -2.3$, $p < .05$), indicating that participants responded to high frequency words more quickly than low frequency words. These results also replicate previous findings (Lewis, Gerhand, and Ellis, 2000; Morrison and Ellis, 2000, as above).

Figure 10. Lexical Frequency on Spanish Word Reaction Time
Main effect of lexical frequency on log reaction time (RT) for Spanish controls and Spanish heritage speakers (SHS) in Spanish. Error bars represent 95% confidence intervals around the mean. Participants responded significantly more quickly to high frequency words compared to low frequency words.



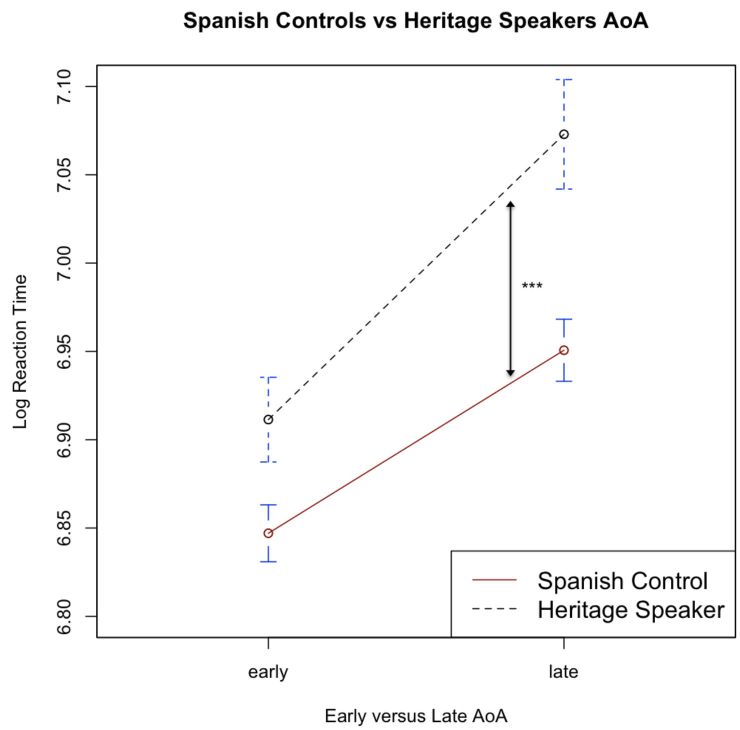
Interactions:

This section will now analyze the interactions between AoA and group (Spanish control versus SHS) and lexical frequency and group. Critically, Spanish controls are L1-dominant in Spanish compared to their SHS peers due to continued language use. If it is the case that extended use is necessary for native-like (L1-dominant like) lexical retrieval, then this predicts faster RTs for L1-dominant Spanish controls compared to SHS (non-dominant L1). Conversely, we argue that language learning mechanisms may actually benefit from early exposure, predicting that SHS will not show RT differences compared to Spanish controls on early-acquired words, but will be more adversely affected (i.e., a RT slow down) on late-acquired words as many of the late-acquired words in this experiment were reported to be acquired after SHS were introduced to English (their dominant L2). If we observe a relatively greater slow down in RT on late-acquired Spanish words, this would suggest that the introduction of English to SHS limited their ability to continue developing Spanish (their L1).

The three-way interaction between AoA, lexical frequency, and group was not significant (β estimate = 0.008, standard error β = 0.03, $\chi^2(1) = .09$, $t = .3$, $p > .05$). Additionally, the interaction between AoA and lexical frequency was not significant (β estimate = 0.03, standard error β = 0.05, $\chi^2(1) = .42$, $t = .6$, $p > .05$). As mentioned above, there have been multiple conflicting results on whether AoA and lexical frequency interact (Zevin and Seidenberg, 2002 compared to Lewis, Gerhand, and Ellis, 2000). However, the hypotheses in this study remain neutral to a lack of interaction between AoA and lexical frequency with respect to Spanish controls versus heritage speakers.

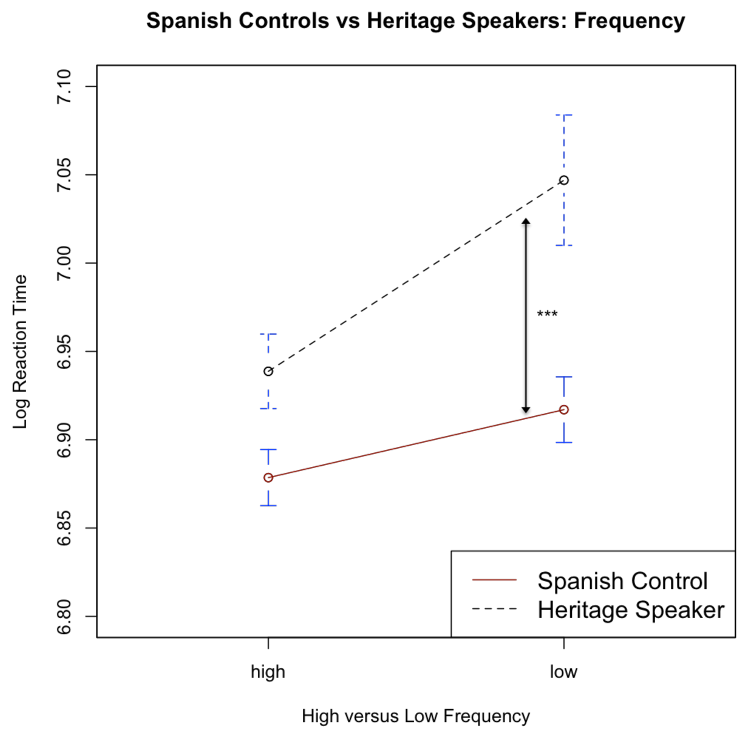
The interaction between AoA and group was significant (β estimate = 0.04, standard error β = 0.02, $\chi^2(1) = 7.23$, $t = 2.6$, $p < .01$) indicating that Spanish controls were less adversely affected (i.e., faster) at identifying late acquired words compared to SHS. This result is displayed in Figure (11) below.

Figure 11. *Interaction between age of acquisition (AoA) and group (Spanish control versus SHS) on log reaction time (RT) for Spanish controls and Spanish heritage speakers (SHS) in English. Error bars represent 95% confidence intervals around the mean. Spanish controls were less adversely affected (responded significantly more quickly, indicated by the non-parallel lines ***) by late-acquired words compared to Spanish controls.*



The interaction between lexical frequency and group was also significant (β estimate = 0.05, standard error β = 0.02, $\chi^2(1) = 8.62$, t = 2.9, p < .01) indicating that Spanish controls were also less adversely affected by low lexical frequency relative to high lexical frequency compared to SHS resulting in a much larger group difference for the low frequency words than for the high frequency words. This result is shown in Figure (12).

*Figure 12. Interaction between lexical frequency (high and low frequency) and group (Spanish control versus SHS) on log reaction time (RT) for Spanish controls and Spanish heritage speakers (SHS) in Spanish. Error bars represent 95% confidence intervals around the mean. Spanish controls were less adversely affected (responded significantly more quickly, indicated by the non-parallel lines ***) by low frequency words compared to Spanish controls.*



3.2.5 Discussion of Spanish Results[23]

The main goal of this section is to determine whether early, but limited, exposure to language is sufficient to prevent any deficits in lexical decision processing at adulthood in speakers whose L1 is dissociated from their dominant language. Specifically, we want to know whether the early exposure of Spanish heritage speakers received as children allowed them to recognize words as quickly as their L1-dominant Spanish speaking peers. From these significant interactions, we see that the early exposure to the heritage language (Spanish) was insufficient to prevent processing deficits in late-acquired or low frequency words. However, it should be noted that despite a slowdown on low frequency and late-acquired words and overall lower

[23] For a side-by-side comparison of all groups and interactions, see the appendix (section 4)

accuracy compared to Spanish controls, SHS still managed to receive a benefit of early-acquired and high frequency words in their L1, despite limited usage as adults. The implications of these results will be discussed at length in Chapter 5 (Discussion of results section 5.1).

Chapter 4. Experiment 3: Word Reading Time (RT) and Duration

In experiment 3, we compared the English and Spanish reading times (RTs) and word durations of Spanish heritage speakers (L2-dominant English) to L1-dominant English and L1-dominant Spanish speakers' RTs, respectively. The goal of this study was to determine whether lexical processing and production mechanisms benefit from early exposure such that listeners can produce words that vary on age of acquisition and lexical frequency in their non-dominant L1 in a similar way that L1-dominant listeners retrieve them. Furthermore, we question whether the delayed, but extended, exposure to the L2 was sufficient to overcome any processing difficulties incurred by later L2 acquisition.

4.1 Methods

4.1.1 Stimuli

The stimuli are the same as described in Experiment 2. While only 64 real words (of either English or Spanish) were used in Experiment 2, all real words in each language (N = 128) generated from the stimuli selection process were used in Experiment 3 (n = 32 in each age of acquisition (early or late) and lexical frequency (high or low) combination such that early-low, early-high, late-low, late-high cells contain the same number of words). Words appearing in only Experiment 3 were labeled as "new" and words appearing in both Experiments 2 and 3 were labeled as "old", in order to assess the affect (if any) of words appearing in Experiment 2 beforehand on word reading time or duration.

4.1.2 Participants

The participants in Experiment 3 are the same participants from Experiment 2 (English controls, Spanish controls, Spanish heritage speakers (SHS), n = 20 in each group). See section 3 for a complete description of these participants.

4.1.3 Procedure

The stimuli were presented individually to participants on a computer monitor in a sound attenuated booth. Stimuli orders were randomized for each participant. Participants were instructed to read each stimulus out loud as it was visually presented as quickly and accurately as possible into a Shure SM81 Condenser Handheld microphone, while clicking “Next” at a fixed location on the monitor to move on to the next stimulus. They were informed that all of the stimuli were real words of either English or Spanish, respectively. The “Next” button was linked to a built-in microphone that marked the boundaries on a separate audio channel between word presentations in order to mark both the onset and offset of stimulus presentations (which allow for reading time and duration measurements to be automatically extracted).

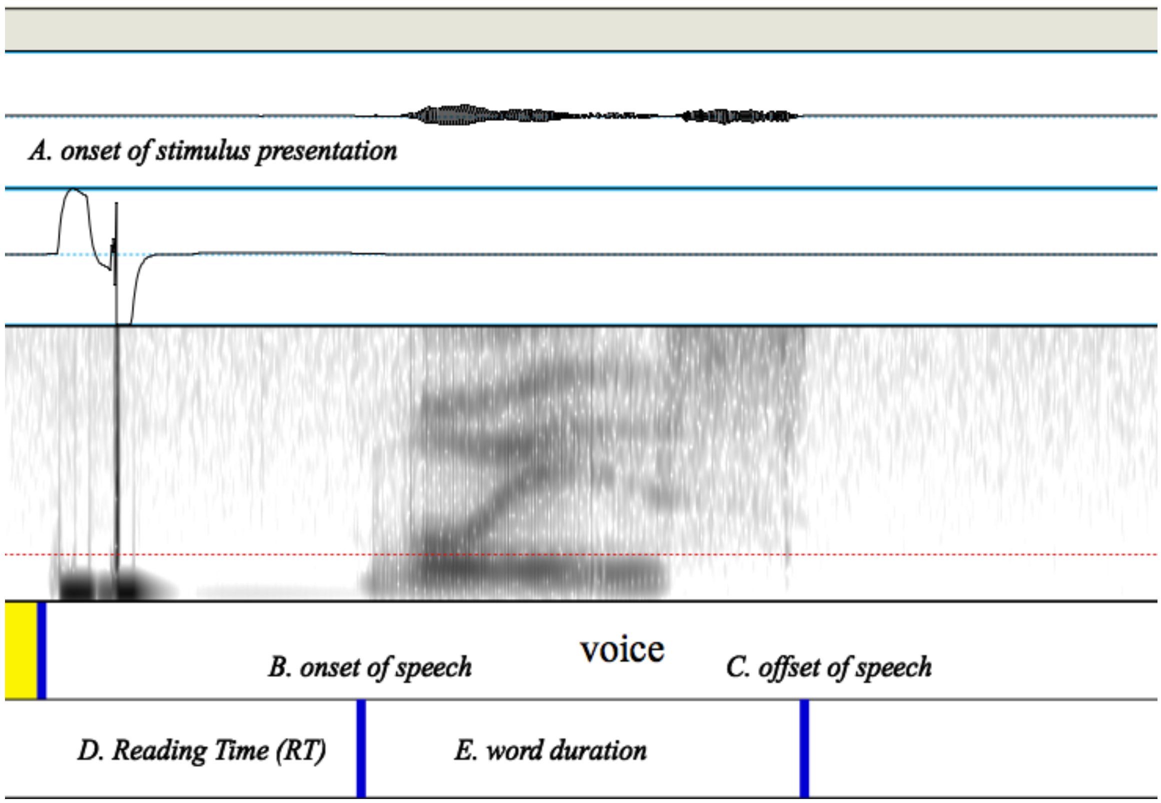
English and Spanish controls (n = 20 in each group) completed Experiment 2 (lexical decision task) first, took a short 5-minute break, and then completed Experiment 3. Spanish Heritage Speakers (SHS, n = 20) completed Experiment 2 first in either English or Spanish and then completed Experiment 3 in the same language after a short break. With respect to SHS, one half of the participants (n = 10) completed both Experiment 2 and 3 in English first, while the other half completed both experiments in Spanish first. SHS and Spanish control participants were paid 10 dollars an hour for their participation in this experiment and experiment 2 combined. English control participants received course credit. Between languages they took an extended break as needed, typically lasting 10-15 minutes. All participants finished Experiment 3 in less than 10 minutes per language, if applicable.

4.2 Results of Experiment 3

4.2.1 Modeling Results

The following section provides the reading time (RT) and word duration results from the above procedure (Experiment 3). Using a script developed in Praat (Boersma and Weenink, 2017), the audio channel that contained the recorded "Next" clicks was used to create an 'onset' and 'offset' boundary grid for the presentation of each stimulus. Then, using a separate Praat script, the onset and offset of each word was obtained by searching within each onset/offset boundary for a specific decibel (dB) threshold (here, at least a relative 40 dB within the interval) to determine the onset and offset of speech (i.e., to extract word duration). The results for each talker were then hand-corrected for errors that the scripts might have missed (lip smacks, coughs, etc.), alignment errors, as well as removing any results that contained reading errors from participants (incorrect words, false starts, etc.). Finally, using a third Praat script, the following measurements were obtained from all hand-corrected TextGrids for each word: onset of presentation of stimulus, onset of speech, reading time (RT) to stimulus (onset of speech – onset of presentation), offset of speech, offset of presentation, and word duration (offset of speech – onset of speech). See Figure (13) below for a visual representation of these measurements.

Figure 13. Sample Speech Segmentation for Experiment 3.
In this figure, the vertical blue bar at (A) is the onset of the presentation of the stimulus "voice" marked using a trigger channel connected to a mouse click. The vertical blue bar at (B) is the onset of speech by the participant, while the vertical blue bar at (C) is the offset of speech. Reading time at interval D (RT) was measured by subtracting time point (B) from (A). Word duration at interval E was measured by subtracting time point (C) from (B).



All data were analyzed using the software R (R version 3.1.0, 2014) using maximal linear mixed effects regressions (LMERs) with random intercepts for participant and lexical items as well as random slopes for the three factors of interest: age of acquisition (AoA: early or late), lexical frequency (high or low) and participant group (heritage speaker or control) by subject (e.g., Barr, Levy, Scheepers and Tily, 2012). Additionally, whether stimuli were "old" or "new" was a single factor (contrast coded, 0.5 for "old" and -0.5 for "new") included in the models to determine what effect, if any, previous exposure to a word in Experiment 2 had on word duration[24] or reading time (RT). The three factors of main interest (AoA, lexical frequency and participant group) were also contrast coded (0.5 for 'high', 'early', and 'control'; -0.5 for 'low', 'late' and 'heritage speaker', respectively) before building any models. Two separate models with two separate dependent variables, reading time (RT) and word duration, were analyzed. Any data more than 2.5 standard deviations[25] above the mean were excluded[26] from the final analyses. The total number of words (out of a possible 2560 words) analyzed for each talker group were: English controls (n = 2431; 95%), Spanish controls (n = 2407; 94%), and SHS (English, n = 2560, 100%; Spanish n = 2494, 97.4%) after incorrect readings, disfluencies and false starts were removed. An independent, phonetically trained rater with knowledge of both English and Spanish was given a random selection of eight participants' recordings (two English controls, two Spanish controls, two heritage speakers in Spanish, two heritage speakers in English) to verify the hand corrections from the Praat script performed by the author. The rater was given no information of how any word varied on AoA or lexical frequency. The results of this verification task confirm that the Praat script and any hand-corrections by the author were highly accurate. On average, there was a 7.37 millisecond difference (the correlation was also high with r = .99 between raters) between reading time measurements and an 8.95 millisecond difference (r = .98) between duration measurements, indicating high degree of reliability. Significance was assessed using the likelihood ratio test in which the degree to which the data are fit by a full model versus a model excluding a factor or interaction of interest is measured. The final (separate) models for both word duration and reading time (RT) in both English and Spanish as the dependent variables converged with a three-way interaction between group, AoA

[24] Reduced duration would be consistent with Baker and Bradlow (2007) and other second mention reduction studies that have shown previously mentioned words to be produced with shorter durations in later discourse

[25] As RT and duration necessarily have a right-tail, only values above this threshold were excluded

[26] Models were run with the original set of data kept in and did not significantly alter results.

and lexical frequency, as well as a three-way interaction between AoA, frequency and whether the stimuli had been previously presented and an interaction between group and previous presentation in addition to all subsequent two-way interactions and main effects. The random intercept structures differed slightly. Initially, both models included fully maximal random effects structures with random slopes for the interaction between AoA and lexical frequency by subject. With respect to word duration, the model that converged included random intercepts for participants and lexical items with random slopes for the interaction between AoA and frequency by subject. However, with respect to RT, this maximal random effects structure failed to converge on the data. A model with random intercepts for participants and lexical items with random slopes for the AoA and frequency by subject with the interaction removed eventually converged on the data. These models are shown below:

Word duration:

```
lmer(duration~AoA.contrast*Frequency.contrast*Group.contrast +
AoA.contrast*Frequency.contrast*Experiment.contrast + Group.contrast*Experiment.contrast
(1+ AoA.contrast*Frequency.contrast|Subject) + (1|Stimulus))
```

Reading Time (RT):

```
lmer(logRT~AoA.contrast*Frequency.contrast*Group.contrast +
AoA.contrast*Frequency.contrast*Experiment.contrast + Group.contrast*Experiment.contrast
(1+ AoA.contrast+Frequency.contrast|Subject) + (1|Stimulus))
```

4.2.2 English Results

Word duration

A. Main Effects

1. Group

There was a significant main effect of group on word duration (β estimate = 0.04, standard error β = 0.02, $\chi^2(1) = 5.01$, $t = 1.93$, $p < .05$) indicating that Spanish heritage speakers (SHS) produced shorter word durations over all compared to English controls (see figure (14) below).

2. Age of Acquisition (AoA)

There was a significant main effect of age of acquisition (AoA) on early acquired words (β estimate = -0.08, standard error β = 0.02, $\chi^2(1) = 19.83$, $t = -4.62$, $p < .05$) indicating that early-acquired words were produced with shorter durations compared to late-acquired words. This result is shown in figure (14) below.

3. Lexical Frequency

There was a significant main effect of lexical frequency on word duration (β estimate = -0.03, standard error β = 0.02, $\chi^2(1) = 4.81$, $t = -2.23$, $p < .05$) indicating that high frequency words were produced more quickly than low frequency words. Main effects are shown below in figure (14) along with the other interactions of interest in this study.

4. Previous Mention in Experiment 2

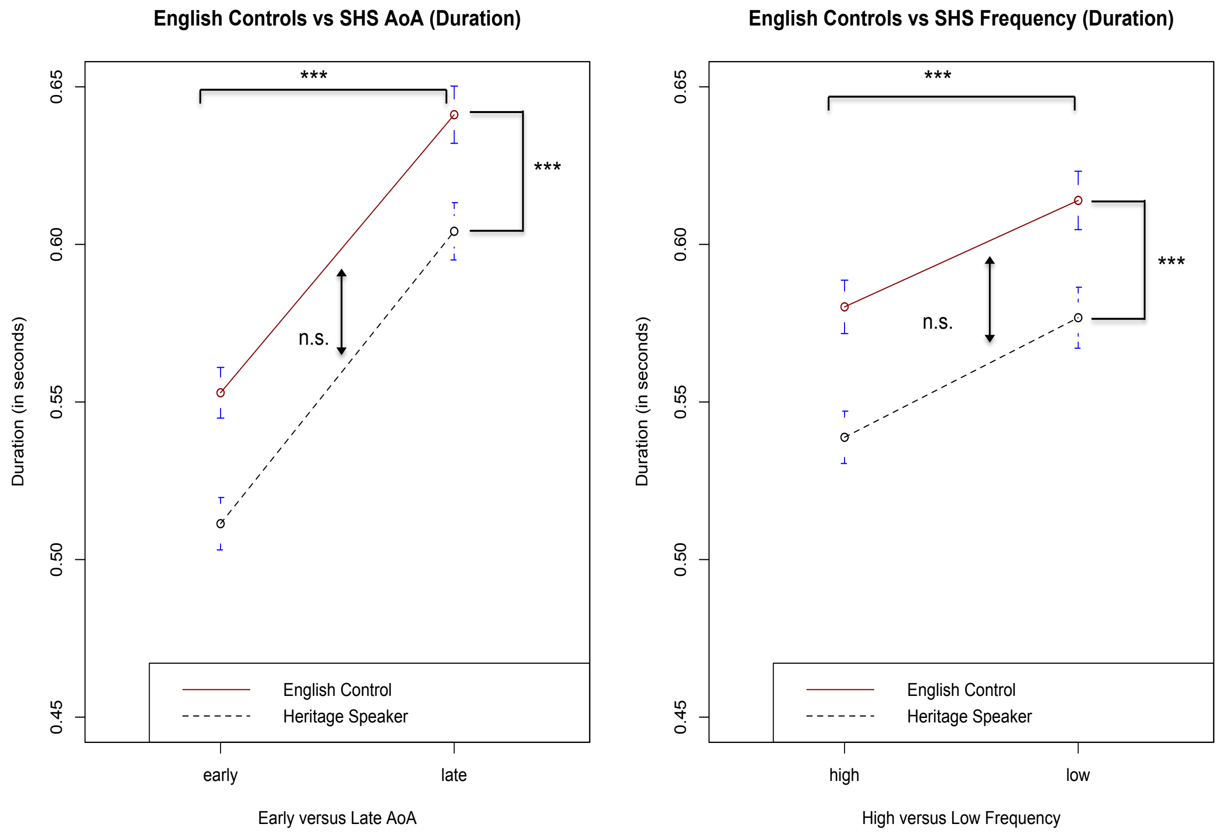
As discussed above, one half of the words present in both Experiment 2 and 3. The results of LMER model comparisons revealed that experiment status ("old" versus "new") was not a significant predictor (i.e., main effect) on word duration (β estimate = -0.02, standard error β = 0.02, $\chi^2(1) = 0.83$, $t = -0.93$, $p > .05$). The three-way interactions between AoA, lexical frequency and experiment status was also not significant (β estimate = -0.11, standard error β = 0.07, $\chi^2(1) = 3.07$, $t = -1.71$, $p > .05$) nor were any of the two-way interactions: experiment status and group (β estimate = -0.001, standard error β = 0.004, $\chi^2(1) = 0.25$, $t = -0.50$, $p > .05$); experiment status and AoA (β estimate = -0.008, standard error β = 0.03, $\chi^2(1) = 0.08$, $t = -0.27$,

$p > .05$); experiment status and lexical frequency (β estimate = 0.07, standard error β = 0.03, $\chi^2(1) = 3.73$, $t = 1.91$, $p > .05$).

B. Interactions

The three-way interaction of group, AoA and SHS was not significant (β estimate = 0.001, standard error β = 0.007, $\chi^2(1) = 0.01$, $t = 0.09$, $p > .05$). The interaction between AoA and lexical frequency was not significant (β estimate = 0.003, standard error β = 0.03, $\chi^2(1) = 0.003$, $t = 0.09$, $p > .05$). The interaction between AoA and group (SHS versus English control) was not significant (β estimate = 0.005, standard error β = 0.005, $\chi^2(1) = 0.82$, $t = 0.90$, $p > .05$), indicating that SHS and English controls word durations were affected by AoA in a similar manner. The interaction between lexical frequency and group was also not significant (β estimate = 0.002, standard error β = 0.004, $\chi^2(1) = 0.01$, $t = 0.45$, $p > .05$), indicating that SHS and English controls were affected by lexical frequency in a similar manner. The results are shown in Figure (14).

Figure 14: Age of Acquisition and Lexical Frequency on English Word Duration
*Both English controls and Spanish Heritage Speakers' (SHS) word durations were equally affected by lexical frequency and age of acquisition (p > .05, indicated by the non-significant parallel lines). On average, SHS produced significantly shorter word durations compared to English controls (p < .05, indicated by *** on right hand side of each figure). Both groups produced shorter word durations on early-acquired words (p < .05, compared to late-acquired words) and high frequency words (p < .05, compared to low frequency words) indicated by *** at the top of each figure. Error bars represent 95% confidence intervals around the mean.*



Reading Time (RT):

A. Main Effects

1. Group

There was no significant main effect of group (β estimate = -0.01, standard error β = 0.02, $\chi^2(1) = 0.14$, $t = -0.87$, $p > .05$) on reading time (RT), indicating that English controls began reading the stimuli as quickly as Spanish heritage speakers (SHS).

2. Age of acquisition (AoA)

There was no significant main effect of age of acquisition (AoA; β estimate = -0.007, standard error β = 0.004, $\chi^2(1) = 3.38$, t = -1.79, p > .05) on reading time (RT), indicating that participants began reading words that were early or late-acquired at approximately the same time as the stimuli were presented to them.

3. Lexical Frequency

There was a significant main effect of lexical frequency (β estimate = -0.01, standard error β = 0.004, $\chi^2(1) = 10.67$, t = -3.33, p < .05) on reading time (RT), indicating that participants were faster to begin reading high frequency words compared to low frequency words. These results are consistent with studies (such as Morrison and Ellis, 1995) that have shown reduced reading times for high frequency words. The independent force of each of these factors will be discussed later in a broader section. These results are shown in figure 15 below.

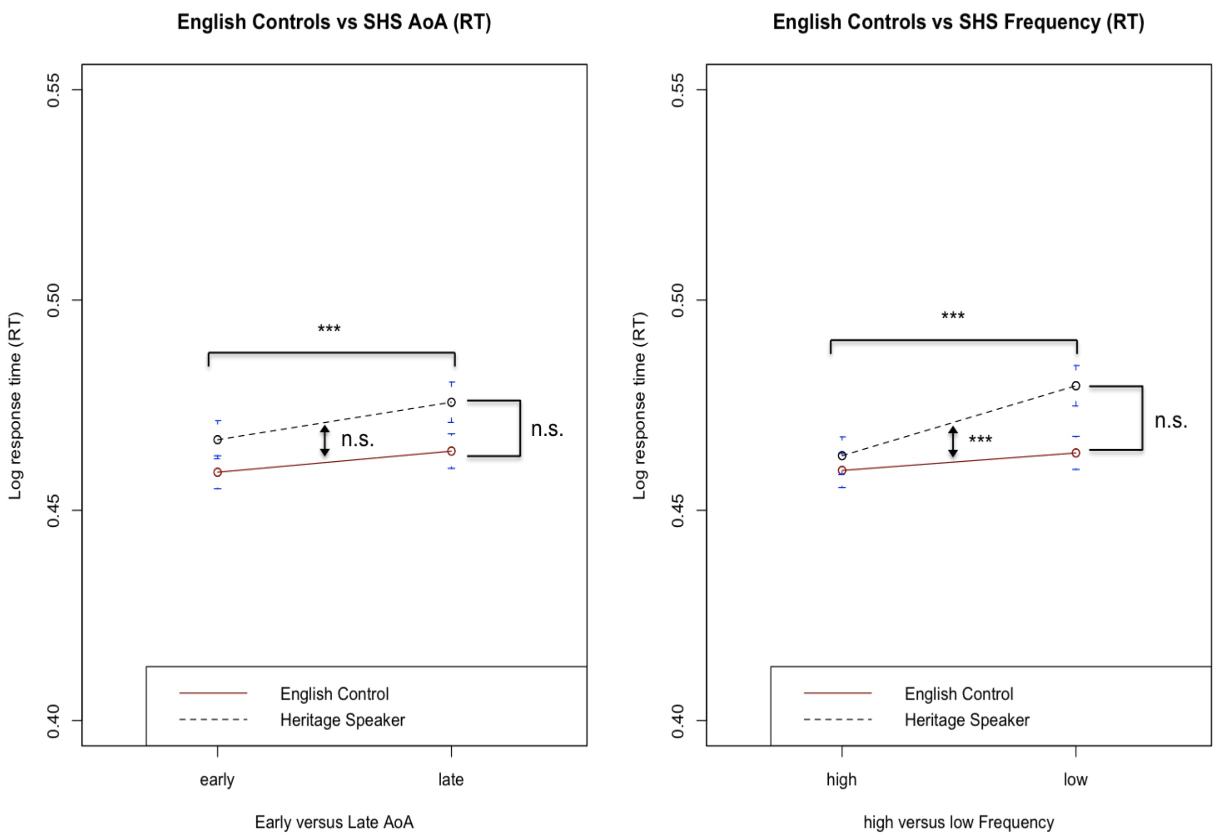
4. Previous Mention in Experiment 2

As discussed above, one half of the words present in both Experiment 2 and 3. The results of LMER model comparisons revealed that experiment status ("old" versus "new") was not a significant predictor (i.e., main effect) on word duration (β estimate = -0.01, standard error β = 0.003, $\chi^2(1) = 1.7$, t = -1.29, p > .05). The three-way interactions between AoA, lexical frequency and experiment status was also not significant (β estimate = 0.002, standard error β = 0.02, $\chi^2(1) = .001$, t = 0.11, p > .05) nor were any of the two-way interactions: experiment status and group (β estimate = 0.003, standard error β = 0.003, $\chi^2(1) = 0.25$, t = -0.50, p > .05); experiment status and AoA (β estimate = 0.01, standard error β = 0.01, $\chi^2(1) = 0.65$, t = 0.79, p > .05); experiment status and lexical frequency (β estimate = 0.01, standard error β = 0.01, $\chi^2(1) = 0.38$, t = 0.60, p > .05).

B. Interactions

The three-way interaction of group, AoA and SHS was not a significant predictor of word RT (β estimate = 0.004, standard error β = 0.01, $\chi^2(1) = 0.44$, $t = 0.66$, $p > .05$). The interaction between AoA and lexical frequency was not significant (β estimate = 0.01, standard error β = 0.01, $\chi^2(1) = 0.37$, $t = 0.61$, $p > .05$). The interaction between AoA and group (SHS versus English control) was not significant (β estimate = 0.01, standard error β = 0.004, $\chi^2(1) = 1.7$, $t = 1.3$, $p > .05$)., indicating that both L1-dominant (English controls) and L2-dominant (SHS) produced similar RTs for both early and late-acquired words. Surprisingly, the interaction between lexical frequency and group was significant (β estimate = 0.01, standard error β = 0.004, $\chi^2(1) = 8.51$, $t = 3.33$, $p < .05$), indicating that SHS were slower to respond to low frequency words compared to English controls. These results are show in Figure (15).

Figure 15. Age of Acquisition and Lexical Frequency on English Reading Time
*Both English controls and Spanish Heritage Speakers' (SHS) reading times (RT) were equally affected by age of acquisition (p > .05, indicated by the non-significant parallel lines). However, SHS were more adversely affected (i.e., slower RT) by low frequency words compared to English controls (p < .05, indicated by the *** parallel lines). Both groups produced shorter RTs on early-acquired words (p < .05, compared to late-acquired words) and high frequency words (p < .05, compared to low frequency words) indicated by *** at the top of each figure. Error bars represent 95% confidence intervals around the mean.*



4.2.3 Discussion of English Results

In this section, we see that in general, SHS (L2-dominant English speakers) were as fast to respond to stimuli over all compared to English controls, yet surprisingly, SHS produced shorter word durations. Additionally, despite a delayed onset of acquisition to English, L2-dominant SHS were not more adversely affected by early-acquired words compared to English controls (we can assume that the acquisition process for late-acquired words would be similar for

SHS and English controls). Interestingly, however, SHS were slower to respond to low frequency English words compared to English controls, but received a similar processing benefit from high frequency words. These results suggest that the frequent exposure to English resulting in (L2) English dominance was sufficient enough to overcome most processing difficulties incurred by the delayed acquisition of their L2. However, SHS were slower in responding to low frequency English words, which suggests that their delayed English acquisition may have impacted at least some part of their overall ability to produce words in English. The implications of these results will be discussed later in a broader context.

4.2.4 Spanish Results

Word Duration

A. Main Effects

1. Group

There was a significant main effect of group (β estimate = -0.05, standard error β = 0.02, $\chi^2(1) = 4.61$, $t = -2.50$, $p < .05$), indicating that Spanish controls produced overall shorter Spanish word durations compared to Spanish heritage speakers (SHS). This result is unsurprising given previous studies that have shown heritage speakers to be slower readers in the non-dominant L1 (Montrul and Foote, 2014; see figure 16 below).

2. Age of Acquisition (AoA)

There was a significant main effect of age of acquisition (AoA) on word duration (β estimate = -0.06, standard error β = 0.02, $\chi^2(1) = 12.81$, $t = -3.49$, $p < .05$), indicating that early-acquired words were produced with shorter overall durations compared to late-acquired words[27] (see figure 16 below).

3. Lexical Frequency

There was a significant main effect of lexical frequency on Spanish word duration (β estimate = -0.04, standard error β = 0.02, $\chi^2(1) = 4.69$, t = -2.10, p < .05), indicating that high frequency words were produced with shorter durations compared to low frequency words. This result is shown in figure 16 below along with the other main effects of interest in this study.

4. Previous Mention in Experiment 2

As discussed above, one half of the words present in both Experiment 2 and 3. The results of LMER model comparisons revealed that experiment status ("old" versus "new") was not a significant predictor (i.e., main effect) on word duration (β estimate = -0.03, standard error β = 0.02, $\chi^2(1) = 2.26$, t = -1.48, p > .05). The three-way interactions between AoA, lexical frequency and experiment status was also not significant (β estimate = -0.05, standard error β = 0.07, $\chi^2(1) = 0.43$, t = -0.64, p > .05) nor were any of the two-way interactions: experiment status and group (β estimate = -0.002, standard error β = 0.004, $\chi^2(1) = 0.25$, t = -0.50, p > .05); experiment status and AoA (β estimate = 0.03, standard error β = 0.04, $\chi^2(1) = 0.85$, t = 0.9, p > .05); experiment status and lexical frequency (β estimate = 0.04, standard error β = 0.04, $\chi^2(1) = 1.29$, t = 1.11, p > .05).

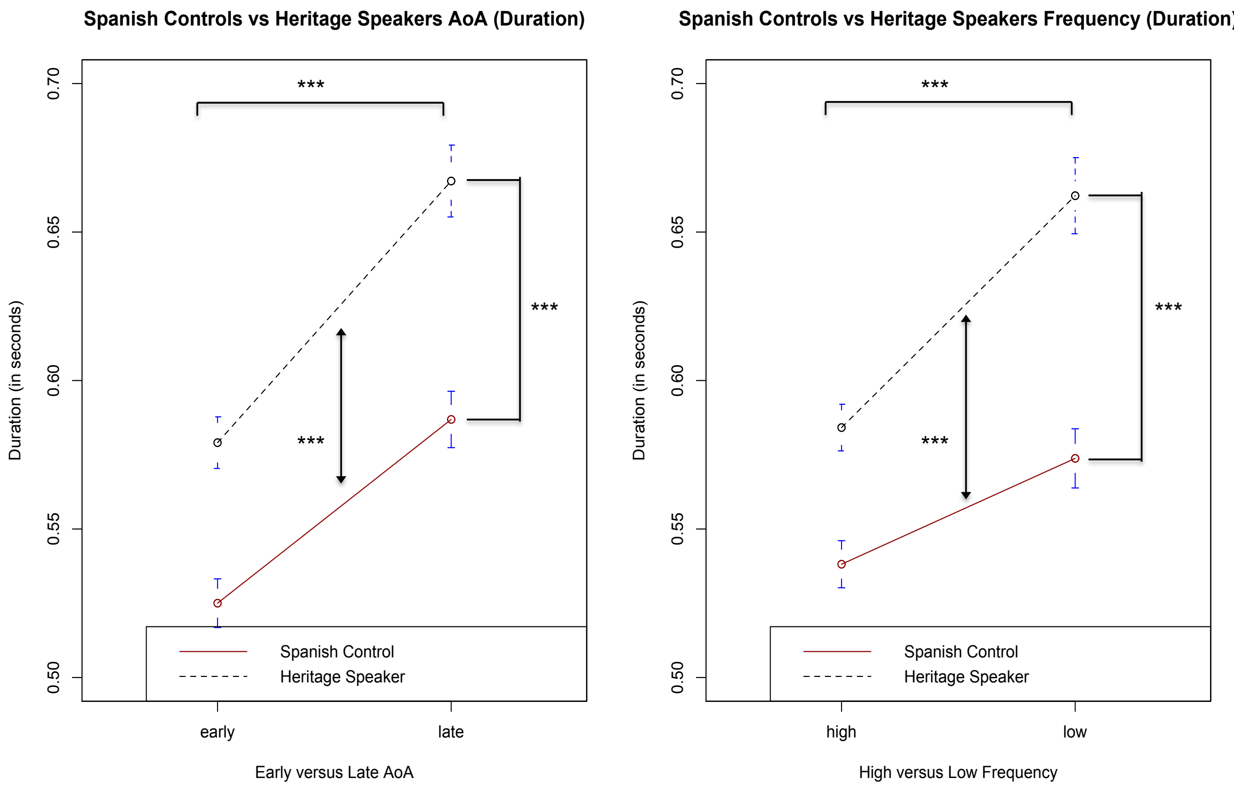
B. Interactions

---

[27] NB: for both English and Spanish word durations and reading times, stimuli were not normalized for phonemic or orthographic length (respectively) as the comparisons of interest in this study are the interactions between AoA/lexical frequency and talker group rather than the main effects of AoA and lexical frequency on words. That is, we are more interested to see if heritage speakers pattern differently compared to controls on these particular features of words (AoA and frequency) rather than how both groups function collapsed together. Furthermore, there was a less than 1-letter and 1-phoneme difference (mean) between high and low frequency words and early and late-acquired words in Spanish; there was approximately a 1.5-letter and 1.5-phoneme difference (mean) between high and low frequency words and early and late acquired words in English.

The three-way interaction between AoA, lexical frequency, and group was not significant (β estimate = -0.002, standard error β = 0.002, $\chi^2(1) = 0.05$, t = -0.23, p > .05) nor was the interaction between AoA and lexical frequency (β estimate = -0.02, standard error β = 0.04, $\chi^2(1) = 0.24$, t = -0.47, p > .05). However, the interaction between AoA (early versus late-acquired) and group (SHS versus Spanish control) was significant (β estimate = 0.01, standard error β = 0.004, $\chi^2(1) = 4.72$, t = 2.21, p < .05), indicating that SHS produced longer word durations on late-acquired words compared to Spanish controls. Additionally, the interaction between lexical frequency (high versus low) and group was significant (β estimate = 0.02, standard error β = 0.01, $\chi^2(1) = 8.83$, t = 3.65, p < .05), indicating that SHS produced longer word durations on low frequency words compared to Spanish controls. These results are shown in Figure (16).

Figure 16. Age of Acquisition and Lexical Frequency on Spanish Word Duration
*Spanish heritage speakers (SHS) were more adversely affecter (i.e., longer word durations) by both late-acquired words and low-frequency words (p < .05, significance indicated by the non-parallel lines ***). On average, SHS produced significantly longer word durations compared to Spanish controls (p < .05, indicated by *** on right hand side of each figure). Both groups produced shorter word durations on early-acquired words (p < .05, compared to late-acquired words) and high frequency words (p < .05, compared to low frequency words) indicated by *** at the top of each figure. Error bars represent 95% confidence intervals around the mean.*



Reading Time (RT)

A. Main Effects

1. Group

There was a significant main effect of group on reading time (RT; β estimate = -0.11, standard error β = 0.04, $\chi^2(1) = 5.77$, t = -3.02, p < .05), indicating that Spanish controls began reading the presented stimuli more quickly than SHS participants. Again, this result is consistent with previous studies (e.g., Montrul and Foote, 2014) that have shown a global slow down in heritage speakers' reading time to stimuli (see figure 17 below).

2. Age of Acquisition

There was a significant main effect of age of acquisition (AoA) on RT (β estimate = -0.04, standard error β = 0.01, $\chi^2(1) = 16.38$, t = -4.38, p < .05), indicating that participants produced shorter RTs for early-acquired words compared to late-acquired words (see figure 17 below).

3. Lexical Frequency

There was a significant main effect of lexical frequency on RT (β estimate = -0.1, standard error β = 0.01, $\chi^2(1) = 26.72$, t = -6.34, p < .05), indicating that participants produced shorter RTs for high frequency words compare to low frequency words. As mentioned earlier, both (high) lexical frequency and (early) AoA reducing overall reading time is consistent with (among many other studies) Morrison and Ellis (1995). This result is shown in figure 17 below along with the other main effects of interest in this study.

4. Previous mention in Experiment 2

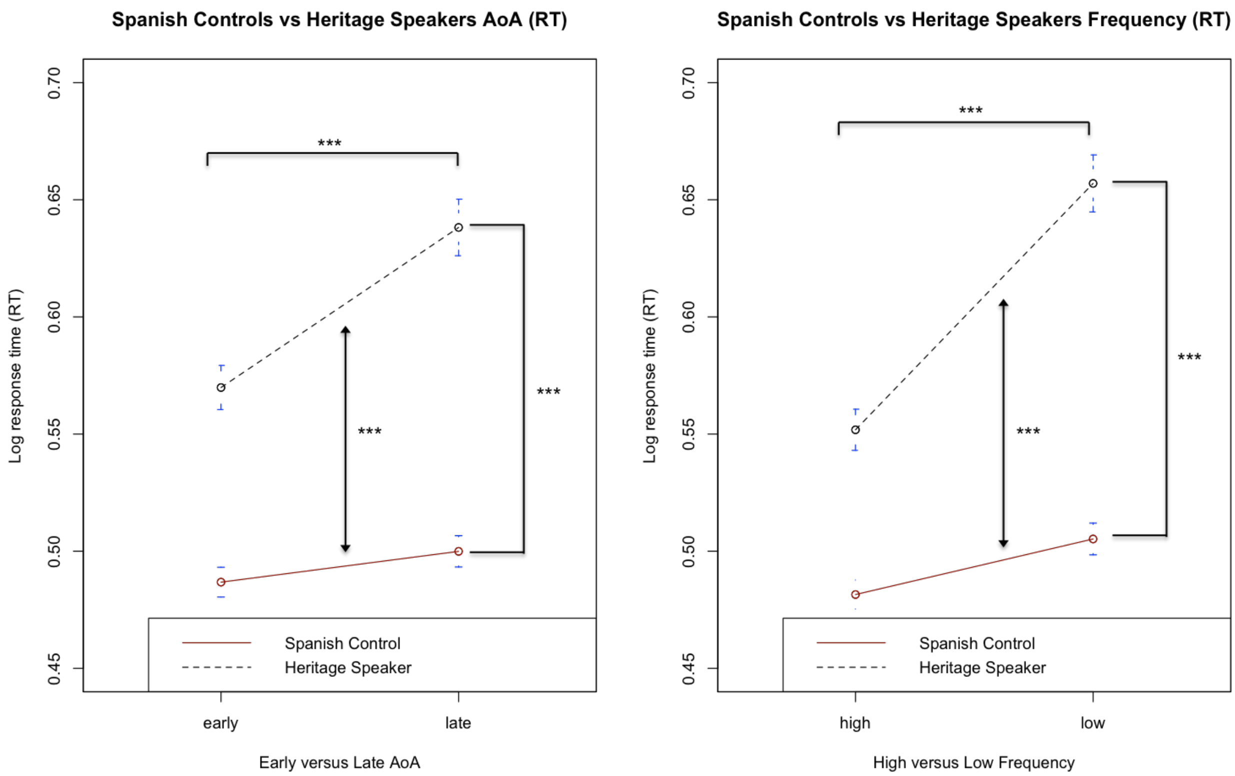
As discussed above, one half of the words present in both Experiment 2 and 3. The results of LMER model comparisons revealed that experiment status ("old" versus "new") was a significant predictor (i.e., main effect) on word duration (β estimate = -0.02, standard error β = 0.01, $\chi^2(1) = 8.87$, t = -3.04, p < .05). This main effect indicates that RTs for words previously heard in Experiment 2 had faster RTs in Experiment 3. However, the three-way interactions between AoA, lexical frequency and experiment status was not significant (β estimate = -0.01, standard error β = 0.03, $\chi^2(1) = 0.14$, t = -0.37, p > .05) nor were any of the two-way interactions: experiment status and group (β estimate = 0.01, standard error β = 0.01, $\chi^2(1) = 2.15$, t = 1.47, p

> .05); experiment status and AoA (β estimate = 0.03, standard error β = 0.02, $\chi^2(1) = 3.14$, t = 1.74, p > .05); experiment status and lexical frequency (β estimate = 0.01, standard error β = 0.02, $\chi^2(1) = 0.66$, t = 0.82, p > .05). These results indicate that despite the general reduction in RT for previously-mentioned test items, the independent effects of group, AoA, and lexical frequency remained unaffected from this second mention.

B. Interactions

The three-way interaction between AoA, lexical frequency and group was not significant (β estimate = -0.01, standard error β = 0.01, $\chi^2(1) = 0.69$, t = -0.84, p > .05) nor was the interaction between AoA and lexical frequency (β estimate = 0.02, standard error β = 0.01, $\chi^2(1) = 1.36$, t = 1.15, p > .05). Critically, however, the interaction between AoA and group was significant (β estimate = 0.06, standard error β = 0.01, $\chi^2(1) = 16.03$, t = 4.41, p < .05), indicating that SHS produced lower RTs when presented with late-acquired words compared to Spanish controls. Additionally, the interaction between lexical frequency and group was also significant (β estimate = 0.08, standard error β = 0.02, $\chi^2(1) = 4.12$, t = 4.98, p < .05), indicating that SHS were slower to respond to low frequency words compared to Spanish controls. These results are shown in Figure (17).

Figure 17. Age of Acquisition and Lexical Frequency on Spanish Reading Time
*Spanish heritage speakers (SHS) were more adversely affecter (i.e., longer RTs) by both late-acquired words and low-frequency words (p < .05, significance indicated by the non-parallel lines ***). On average, SHS produced significantly longer RTs compared to Spanish controls (p < .05, indicated by *** on right hand side of each figure). Both groups produced shorter RTs on early-acquired words (p < .05, compared to late-acquired words) and high frequency words (p < .05, compared to low frequency words) indicated by *** at the top of each figure. Error bars represent 95% confidence intervals around the mean.*



4.2.5 Discussion of Spanish Results

In this section, we see that in general, SHS were slower to respond to stimuli and produced longer word durations over all compared to Spanish controls. This result itself is not surprising—these SHS were not fluent readers of Spanish and as such, a general slow down in reading rates was expected (see Montrul and Foote, 2014). Interestingly, SHS benefitted from early-acquired and high frequency words in their (non-dominant) L1. This result is consistent with other studies on heritage speakers (e.g., Knightly, Jun, Oh, and Au, 2003) that have shown the benefit of early exposure to language learning. However, SHS were more adversely affected (i.e., are the slowest to respond and produced the longest durations) by words that are either late-acquired or low frequency in their non-dominant L1 compared to L1-dominant Spanish controls. These slow downs suggest that the early, but limited and interrupted, exposure SHS received to Spanish was insufficient to overcome the processing difficulties of this particular set of words.

Chapter 5. Concluding Discussion

5.1 Summary of Results

In the current study we questioned the extent to which early exposure can lead to native-like performance on language-based tasks independent of extended exposure over time. Which level(s) of linguistic representation can benefit from limited, early exposure typically seen in infancy and young childhood without continued exposure throughout middle childhood and adolescence? By examining switched-dominance bilinguals (L2-dominant speakers or heritage speakers), we are able to separate the confounding factors of early age-of-acquisition from language dominance that is typically seen in monolingual and bilingual studies. We hypothesized that speech and vocabulary learning mechanisms, unlike morphosyntax, are equipped to capitalize on the salient linguistic features of a small amount of isolated language exposure. If this type of linguistic exposure is early in the language learning process, we predict that heritage speakers will be able to benefit from broad features of speech and vocabulary learning even in their non-dominant L1. However, this also predicts that less salient features that still contribute to a native speaker's overall linguistic system will suffer due to a lack of extended use. That is, the non-dominant L1 will still exhibit some vulnerabilities in speech (e.g., not as intelligible as L1-dominant speakers in all listening situations) and vocabulary (e.g., difficulty processing low frequency or late-acquired words) learning at adulthood. Conversely, if the exposure is sufficiently extended throughout adolescence and adulthood, then we predict that language-learning mechanisms will be able to develop a fully robust linguistic system, which appears native-like despite being acquired later. Three experiments, with three different experimental paradigms, were conducted to help answer the above question on language acquisition.

In the first experiment, the speech intelligibility scores (percent words correctly recognized) of Spanish heritage speakers' (SHS; L2-dominant English) English and Spanish sentence productions were compared to the sentence productions of L1-dominant speakers of English and Spanish, respectively. Native (L1-dominant) listeners were presented with these sentences either embedded in a relatively easy level of noise (-4 dB signal-to-noise ratio, SNR) or in a relatively difficult level of noise (-8 dB SNR) to highlight potentially subtle vulnerabilities in the SHS' speech production not necessarily seen in favorable speech conditions. In the easier listening condition (-4 dB SNR), both the English and Spanish speech intelligibility scores of the SHS participants were within the same range as L1-dominant talkers in both languages. This result supports the claim that speech-learning mechanisms capitalize on early exposure as heritage speakers' (non-dominant) L1 was as intelligible in the easier listening condition (-4 dB SNR) as L1-dominant talkers' L1, consistent with Au et al (2002) and Oh et al (2003). The results of this experiment also support the claim that language learning mechanisms capitalize on extended exposure (resulting in language dominance) as the range of speech intelligibility scores for SHS in the dominant L2 (English) was similar to L1-dominant English talkers in both listening conditions (easy and difficult), despite English being acquired significantly later for the SHS participants. However, there is also evidence that the early, but limited, exposure to the L1 was insufficient to fully protect against degradation in speech learning. In the harder listening condition (-8 dB SNR), SHS' Spanish (non-dominant L1) speech intelligibility scores were significantly worse than L1-dominant Spanish speakers' scores. Early exposure alone cannot account for this data. The results of experiment 1 suggest that despite some benefit to speech intelligibility in the non-dominant L1, early exposure was insufficient to result in native-like speech productions from heritage speakers that would allow native listeners to overcome the processing difficulties of their speech in harder listening condition. These results also suggest that speech learning may still require some extended use over time in order to fully develop native-like levels of performance.

In the second study, we probed the extent to which vocabulary learning mechanisms capitalized on early exposure by comparing the lexical decision reaction times (RT) of SHS compared to L1-dominant English and L1-dominant Spanish listeners to both English and Spanish words, respectively. The words in this study varied on age of acquisition (AoA) and lexical frequency, which are known factors that affect L1-dominant listeners (e.g., Dirix and Dupoux, 2017a, Gerhand and Barry, 1999;). L1-dominant listeners have been shown to respond more quickly to early-acquired words (compared to late-acquired words) and high frequency words (compared to low frequency words). Results showed that with respect to their (L2) dominant language of English, SHS patterned similarly to L1-dominant English listeners. Both SHS and L1-dominant English listeners benefitted from early-acquired and high frequency words as demonstrated by faster reaction times on these words compared to late-acquired and low frequency words, respectively. Furthermore, SHS were not more adversely affected by late-acquired or low frequency words in English compared to L1-dominant English listeners and were as accurate at correctly identifying real English words as L1-dominant English listeners[28]. Despite the delayed acquisition of English, these results suggest that vocabulary learning mechanisms were able to capitalize on the extended use that resulted in language dominance in order to provide the same native-like benefits to lexical retrieval as a native (L1-dominant) speaker. Conversely, in their non-dominant L1 of Spanish, SHS were significantly slower overall and less accurate at correctly identifying real Spanish words compared to L1-dominant Spanish

[28] SHS were significantly less accurate at identifying non-words of English (i.e., failure to correctly say ‘no’ when presented with an audio clip like ‘murg’). The implications of this result will be discussed below.

listeners. Furthermore, while SHS received a benefit of early-acquired and high frequency words, they were more adversely affected (i.e., slower RTs) by late-acquired and low frequency words compared to L1-dominant Spanish listeners. This result is also consistent with Gollan et al (2013) and Montrul and Foote (2014) who have shown that SHS are more adversely affected by late-acquired and/or low frequency words. We interpret this result as evidence that with respect to vocabulary learning, the benefit that early exposure provided to language learning mechanisms is still limited to a small set of salient features of words (e.g., high lexical frequency, early-acquired words) compared to the overall vocabulary seen in an L1-dominant talker.

In the third experiment, we measured the duration and reading time of words produced by SHS speakers in both Spanish and English and compared them to L1-dominant Spanish and L1-dominant English controls, respectively. The words in this experiment also varied on AoA and lexical frequency. The results of this experiment mirrored experiment 2. SHS speakers were as quick to begin reading words (reading time, RT) in English compared to L1-dominant English controls. The only vulnerability observed in the SHS English productions was that SHS speakers were slower to begin reading low frequency words in English. This is again consistent with Gollan et al (2008, 2013) who have shown that bilingual speakers are more adversely affected by low frequency words in either language. Surprisingly, SHS produced overall shorter word durations in English compared to L1-dominant English controls[29]. However, SHS word durations were produced nearly identically to L1-dominant English controls with no interaction between word duration and AoA or lexical frequency (i.e., both SHS and L1-dominant English controls produced, to the same extent, shorter word durations for early-acquired and high frequency

[29] To be discussed below in section 5.6

words compared to late-acquired and low frequency, respectively). The English RT and duration results suggest that the extended exposure to the L2 was sufficient to overcome the delayed onset in English acquisition, resulting in word production sensitivity to both AoA and lexical frequency in English. Conversely, in their non-dominant L1 Spanish, SHS speakers were significantly slower to begin reading Spanish words and produced significantly longer word durations compared to L1-dominant Spanish controls. Furthermore, SHS speakers were more adversely affected by late-acquired and low frequency words (i.e., slower reading times, longer word durations) compared to L1-dominant Spanish controls. However, we observed a benefit of early-acquired and high frequency words for SHS speakers. That is, despite an overall slow-down in reading time and word durations, SHS managed to benefit from the same features (early AoA and high lexical frequency) as L1-dominant Spanish controls (i.e., faster reading times, shorter word durations). Again, as in experiments 1 and 2, the mixed successes and failures in SHS' non-dominant L1 suggest that speech and vocabulary learning mechanisms are able to capitalize on early exposure, but that there is a certain amount of extended use necessary in order to achieve fully native-like representations.

5.2 Early Acquisition Versus Extended Use

These combined results contribute new information to the debate over whether age of acquisition and lexical frequency act as independent forces on lexical processing. If it were the case that early exposure per se had no effect on lexical processing independently from the effect of extended use over time (i.e., early acquisition is just extended use over time as claimed in Zevin and Seidenberg, 2002), then we would not expect to see heritage speakers show any independent processing benefit for early-acquired words over late-acquired words in the non-dominant L1 (Spanish), and only expect lexical frequency as an independent main effect on SHS word retrieval and production in Spanish. The results of experiment 2 and 3 clearly do not support this claim. Heritage speakers, in fact, did benefit from early-acquired Spanish words independent of lexical frequency (i.e., no statistical interaction). Thus, the SHS in this study are consistent with Gerhand and Barry (1999) who have shown that AoA and lexical frequency are independent forces on the lexicon of L1-dominant speakers.

How might these AoA and lexical frequency features be maintained over time in any language user? An exemplar-based model (e.g., Johnson, 1997; Pierrehumbert, 2001; 2006) of lexical representations may explain why high frequency words are retrieved faster and produced more quickly. The episodic trace models that have followed (e.g., Goldinger, 1996) have shown that listeners store non-linguistic features of words (e.g., indexical information about the talker, background noise, etc.) for some time after being exposed to them. Having observed significant main effects of lexical frequency (i.e., number of occurrences of a word) on word processing, it must be the case that the frequency of a word is also stored for some extended period of time in a language user. The number of occurrences of a particular word strengthening its overall exemplar representation can easily explain this effect of lexical frequency on word processing. However, these models are unable to predict the robust effect of AoA replicated in this study and observed in many others (e.g., Belke et al, 2005; Carroll and White, 1973; Cortese and Khanna, 2007). While not a controversial statement given the wide range of research on this topic, it is still unclear how long the effect of AoA persists in word retrieval throughout a language user's experience. Monolinguals are unable to provide a clear answer for the length over time to which these features are stored, as they confound timing (early acquisition) and amount (extended use). By producing and perceiving early-acquired words more quickly than late-acquired words in the non-dominant L1, the results from heritage speakers in this study suggest that the linguistic system is highly sensitive to early input, causing words acquired later to compete with early-acquired words for lexical processing resources. That is, it could be the case that early-acquired words imprint on the blank linguistic system (during infancy), defining the exemplar space such that later-acquired words must be accommodated at a processing cost. As the effect is seen in both the L1 and the L2 of HS (and the L2 of other bilinguals, cf. Izura and Ellis, 2002), it must be the case that, at least in part, these exemplar spaces contain some separation and/or distinction based on the language being used. Without such a distinction, the benefit (e.g., faster RT) to lexical processing seen in both the L1 and L2 for early-acquired words could not be possible, as the words from the L1 would have already defined the total space dedicated to lexical processing. As their heritage language clearly isolates early exposure, HS are one such group that shed light on the way AoA may be incorporated into exemplar models of lexical retrieval.

Another question that arises from then is how language learning mechanisms successfully capitalize on such little input. Modeling and behavioral studies that have posed this question have shown that in addition to being early, this relatively small amount of input must also be first for successful language development. Computer models of neural networks (e.g., Altmann and Dienes, 1999; Elman, 1993) have shown that language acquisition simulations with totally blank models (parallel to a new born being exposed to a language at home with no prior linguistic experience) can rapidly acquire linguistic features with only a few data points. That is, given a neural network with no prior exposure to language, very little input is required at first for the network to abstract over particular salient linguistic features, resulting in a native-like representation of language. This modeling work is also consistent with behavioral research such as Vlach (2014) who has shown that toddlers also can rapidly learn new vocabulary items by selecting the most salient shared feature across various tokens. In Vlach (2014), toddlers given a novel item and its corresponding word (a new toy, etc. that has never been seen before with a label that has never been heard before by the child) were able to successfully acquire the new word after only a few exposures to that item that varied on non-salient features (i.e., color) but remained constant on one salient feature (i.e., unique shape). Combined, these results have shown that language-learning mechanisms seem to be able to rapidly sort out and filter which features are necessary for successful acquisition from a limited amount of linguistic exposure (assuming the exposure was the first of its kind encountered by the language learner). While able to account for their (partially) successful non-dominant L1 acquisition, how do heritage speakers then acquire their second language in such a native-like way? It must be the case that the extended use of heritage speakers' L2, which results in L2 language dominance, is necessary for successful language acquisition at later stages in the language learning process. If true, this claim would explain how both the non-dominant L1 Spanish of SHS benefitted from features that supported successful language communication (e.g., lower levels of noise, early-acquired words, high lexical frequency words) and the dominant L2 English was not adversely affected by any stresses in linguistic processing (e.g., higher levels of noise, late-acquired words, low lexical frequency words). However, the amount of exposure that would be minimally necessary for successful late acquisition is still unclear. Heritage speakers, having demonstrated successful acquisition of a later (but highly utilized) language, serve as an empirical testing ground to measure the amount of exposure necessary at later stages of language acquisition. A future study in which the variation of onset and amount of L2 exposure were more carefully controlled would be able to help refine our understanding of the processes involved in switched-dominance bilingualism.

5.3 Defining a Native Speaker

Combined, the results of these three experiments add a nuanced view to defining a native speaker. Much of the work on modeling bilingualism uses the L1 (first-acquired) language as a baseline to predict how bilingual speakers will behave in the L2. For example, the Perceptual Assimilation Model (PAM; e.g., Best, 1994; Best, McRoberts, & Goodell, 2001) and the Speech Learning Model (SLM; e.g., Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004) argue that at the segment level, the phonological structure of the L1 will predict and/or influence the success (or lack thereof) of L2 production. While these models are at the phonetic and/or phonological level, we can still predict using these same principles how other levels of linguistic representation will be affected in a second language. They leave little to no room for the unique situation that L2-dominant speakers (i.e., heritage speakers) provide. It has been shown in many studies (e.g., Parlato et al, 2011; Polinsky, 2008) that the linguistic structure(s) of a bilingual's dominant L2 may predict linguistic performance rather than the non-dominant L1. Such a result was seen in Polinsky (2008) when Russian Heritage Speakers (RHS) failed to produce accurate Russian case markers due to the influence of English morphological structure, which lacks case marking on nouns. The claim that the structure of a speaker's dominant L2 may predict linguistic performance does not undermine the predictions laid out by PAM or the SLM or any other model of bilingualism—the goal here is to add to the existing models to incorporate speakers for whom an L1-L2 distinction is actually insufficient to predict behavior in both languages. As such, language dominance clearly predicts a wide variety of linguistic outcomes and should be incorporated into any working model of bilingualism, in addition to focusing on order.

With respect to neural processing, there is also some evidence that a dominant L2 can function in a native-like (here, indistinguishable from an L1-dominant language user) capacity. In Pallier, Dehaene, Poline, B., LeBihan, Argenti, Dupoux, and Mehler (2003), Korean adoptees living in France with no recollection of Korean demonstrated neural responses to Korean stimuli that were no different from their responses to other foreign languages. In this respect, these L2-dominant French speakers were not different from L1-dominant French speakers. However, L1-dominant French speakers showed more neural activation to French compared to the L2-dominant (early but complete attrition of L1 Korean) French speakers. This result would suggest that the sole, extended exposure to French was enough to "tune out" other languages similar to L1-dominant speakers but that the delayed exposure to French also affected general neural activation. However, in this same study, behaviorally, L1-dominant and L2-dominant French (non-dominant L1 Korean) speakers both failed to identify Korean sentences or word translations. While it may be the case that some neural activation between heritage speakers and L1-dominant speakers of the same language is slightly different, the resulting behavioral output is largely unaffected, suggesting that at least with respect to the communicative function of language, heritage speakers are native-like in their dominant L2. Even more broadly, this suggests that the neural associations of language are not static and can adapt as long as the exposure of a particular language is at a sufficient level such that language-learning mechanisms can eventually capitalize on it (e.g., Perani and Abutalebi, 2005; Werker and Tees, 2005). As there is ample evidence that extended exposure results in performance measures similar to L1-dominant speakers, one could define native-like language processing at any level as an idealized standard of performance from monolingual speakers rather than being defined by the order in which the language was acquired. This definition allows us to claim that, in fact, heritage speakers are native speakers of their dominant second language as we define native-like performance to be similar to linguistic output that is observed in L1-dominant and/or monolingual speakers. This is not to say that a heritage speaker is exactly parallel to a monolingual speaker in their dominant L2. In fact, there is nothing to suggest that heritage speakers would not pattern like other bilingual speakers (e.g., SHS demonstrated a frequency lag effect in this study on low frequency English words similar to other bilinguals in Gollan et al, 2013). However, the focus is that, behaviorally, we have ample evidence from this study and a number of other studies (e.g., Albirini et al 2011; Parlato et al, 2010; etc.) that the dominant L2 of heritage speakers can function like a dominant L1 speaker strengthening the idea that the dominant L2 of a heritage speaker is, for all intents and purposes, their "native" language.

The claim that heritage speakers are native speakers of their dominant L2 then raises an interesting question about the non-dominant L1. Is their L1 also a native language or is it the case that their heritage language is closer to an L2 (for a complete review, see Montrul, 2012)? At the behavioral level, we have some evidence that the non-dominant L1 is at least partially distinct from late L2 learners. For example, as mentioned before Blasingame and Bradlow (in prep) found that despite some vulnerability in their non-dominant L1 of Spanish compared to L1-dominant Spanish participants, SHS were still able to outperform late L2 learners of Spanish (L1-dominant English). This result was demonstrated by overall higher word identification accuracy compared to L2 Spanish learners and a benefit of semantic predictability (semantically predictable sentence frames yielded higher sentence final word identification scores) in Spanish where late L2 Spanish learners received no such benefit. Montrul and Foote (2014) also showed at least a partial difference between SHS and late L2 learners such that in a lexical decision task SHS were faster to respond to real words of Spanish compared to late L2 learners. Zamora (2017) also found that SHS (non-dominant L1 Spanish) processed verbal morphology in a distinct manner from late L2 Spanish learners, employing specifically distinct linguistic strategies (e.g., grammaticality judgments) compared to late L2 learners (e.g., recall of explicit instruction of verb conjugation tables) when given particular morpho-syntactic tasks. In this sense, the heritage speaker's non-dominant L1 was still processed differently from a late L2 learner. However, in these three cases, as in the current study, the heritage language was not produced and/or perceived in exactly the same way as an L1-dominant speaker, suggesting that native-like representations of language could be measured on a gradient and/or spectrum rather than bound by some categorical membership (i.e., either a speaker is or is not a native speaker of a language). Heritage speakers make an ideal testing population to explore this potentially nuanced variation within language processing.

5.4 Processing the Non-dominant L1

These observed differences of heritage speakers at various levels of linguistic structure (e.g., speech versus morphosyntax learning) may be explained by a difference in how the linguistic information is processed in a particular experimental paradigm. As such, it may be the case that processing differences (that result from different exposure patterns) can account for the varying successes and failures observed in heritage speaker studies. As mentioned earlier, Zamora (2017) has shown that Spanish heritage speakers (non-dominant L1 Spanish) used distinct strategies compared to late L2 learners (non-dominant L2 Spanish) when processing verbal morphology in Spanish. Specifically, in this study SHS relied on implicit grammaticality judgments based on prior experience with Spanish (in a similar way to an L1-dominant speaker) compared to late L2 learners who relied on more explicit strategies such as inflectional morphology tables learned in a classroom setting. The reason some performance measurements of heritage speakers may be lower (or simply different) in the non-dominant L1 compared to an L1-dominant speaker could be due to processing differences based on the task involved. This result could explain why Montrul and Foote (2014) found a relatively weak benefit of age of acquisition in Spanish lexical translation tasks in SHS—the task itself relied on explicit linguistic processing that heritage speakers are not accustomed to in the heritage language. An explicit versus non-explicit processing account may then explain the variation (i.e., degree of difference from native-like performance) seen between speech and morphosyntax research with the former tasks being measured in a less explicit way compared to the latter being more explicitly probed. That is, it may be the case that morphosyntax experimental paradigms collected data more explicitly in the form of some overt grammatical question compared to speech production and perception paradigms, which utilized less explicit methods (e.g., vowel spaces or VOTs of a heritage speaker).

The case marker *a* in Spanish is one such example of a morpho-syntactic task that may be affected by the availability of explicit versus non-explicit processing. In Montrul (2010b), SHS participants[30] (L2 dominant English) were explicitly asked to judge the grammaticality (an offline method) of Spanish sentences containing *a* case markers and failed to demonstrate native-like (L1-dominant) levels of accuracy and/or sensitivity to the presence or absence of the case marker. Jegerski (2016) showed that, in fact, when instructed to perform self-paced reading tasks (an online method) of Spanish sentences with and without this case marker, heritage speakers were sensitive to its necessary presence or absence.. That is, when given an online (i.e., less explicit) task in which the case marker under question was not explicitly questioned, heritage speakers were able to read grammatical sentences in Spanish more quickly when the case marker was correctly present or absent compared to when its presence or absence produced an ungrammatical sentence. Jegerski (2016) argues that the reason a heritage speaker may not be

[30] Jegerski (2016) also examined late L2 learners who showed similar sensitive to case marking, but critically, they were explicitly taught this marker in classroom settings.

sensitive to a particular feature of the heritage language is that offline measures may not be an available processing strategy in the early, but non-dominant language compared to a late learner who may have a wide variety of explicit processing strategies gained from classroom instruction.

The current study has also employed both on- and offline tasks such that these results may be accounted for in part by this processing distinction. Online processing occurred in experiment 1 via speech intelligibility scores[31] while offline processing occurred in experiments 2 and 3 via a lexical decision task and single word-reading task. Consistent with other speech processing studies that employ more online methods to collect data (e.g., Au et al 2002; Chang et al, 2011, etc.) we found that heritage speakers, in at least one listening condition, were able to achieve speech intelligibility scores in their non-dominant L1 Spanish on par with L1-dominant talkers. Conversely, in experiments 2 and 3, when heritage speakers were asked to judge the lexical status of a word (experiment 2) or read a word out loud when presented to them under timed conditions (experiment 3) in their non-dominant L1, they were unable to reach an L1-dominant range of performance for any measurement (i.e., accuracy, reaction time, reading time, and word duration). These differences could also be explained due to a lack of native-like literacy in the non-dominant L1 observed more generally in heritage speakers (e.g., Polinsky and Kagan, 2008; Montrul, 2010a). In other words, an explicit versus non-explicit processing account could simply be the result of a lack of adult literacy in the non-dominant L1. One way to remove this confound, would be to compare the ability of L1-dominant, but not literate, Spanish speakers to perceive and produce words against SHS in order to determine the role of literacy in lexical processing. That is not to say that the age of acquisition and lexical frequency of words were not

[31] Recall that in experiment 1, the heritage speakers' speech in both English and Spanish was presented to and analyzed by L1-dominant listeners in their respective L1 under various levels of background noise. The L1-dominant listeners, who served merely as a means of gathering speech intelligibility scores for talkers, were explicitly told to write down everything they heard and were made aware of the presence of background noise. Heritage speakers had no knowledge of this manipulation prior to or during the recording of their speech. Therefore, we consider HS speech intelligibility scores to be the result of an online processing task.

been shown to affect heritage speakers' ability to perceive and produce them. However, the overall differences seen in heritage speakers relatively poor performance in their non-dominant L1 (e.g., lower accuracy, slower response times) compared to dominant L1 controls may also be the result of increased processing difficulties due to both literacy issues and more offline measures in word retrieval and production experimental paradigms. As such, the results of this study suggest that more work is needed in order to determine the extent to which offline processing difficulties arise in a non-dominant L1.

5.5 Pedagogical Implications

Another issue that arises in switched-dominance bilingual research is with respect to language education resources for heritage speakers. We ask that if it is the case that heritage speakers at times fail to produce native-like language in their non-dominant L1 and yet are still distinct from late L2 learners, should they still be given the same type of language instruction as late L2 learners? This question has gained wide attention (e.g., Gathercole, 2017; Montrul, 2010a; Montrul 2012; Peyton, Ranard, and McGinni, 2001; Polinsky and Kagan, 2007) and the general focus has been on morphosyntax, as heritage speakers have been shown to pattern more similarly to late L2 learners with respect to this type of learning (See the introduction for a complete review). However, as Haastrup and Henriksen (2001) point out, this focus is limiting in our general understanding of (second) language acquisition. More attention needs to be given at the vocabulary level in both heritage speaker and second language learner education. While the focus on speech learning for heritage speakers may not be critical, due to their relatively good speech perception and production skills in the heritage language (e.g., Au et al, 2002; Oh et al, 2003, etc.), other areas such as reading (e.g., Montrul, 2010a) and word processing (e.g., Gollan et al 2013) may require more attention than previously thought. The results of the current study suggest that heritage speakers would most benefit from increased exposure to low frequency and/or late-acquired items in Spanish to overcome the processing difficulties of these relatively rare forms (i.e., the negative effects of the frequency-lag seen in their heritage language). The discussion above on the difficulties associated with explicit linguistic processing may also guide education policy. Although heritage speakers have been shown in this study (and others) to produce L1-dominant-like speech in their non-dominant L1, they still exhibit some vulnerabilities, especially when the linguistic system is stressed (i.e., their speech is subject to degradation). Additionally, as tasks in which heritage speakers must explicitly analyze linguistic features (e.g., lexical status) have been shown to come at a processing cost (e.g., Jegerski, 2016), instruction in which HS are taught strategies to explicitly analyze their heritage language would also be useful in a classroom setting. While not specifically designed for it, the current study may be able to provide more general insight on the importance of more explicit instruction for all levels of linguistic structure (including speech learning) and strategies that heritage speakers can utilize when faced with explicit processing tasks in their non-dominant L1.

5.6 Limitations and Future Directions

One curious result that is difficult to explain is that in both Experiments 2 and 3, SHS participants responded more quickly (faster reaction time) and produced consistently shorter words (overall shorter duration) in English compared to L1-dominant English speakers. These results were not driven by AoA or lexical frequency (i.e., no interaction), so they must stem from some other linguistic or general cognitive component. While outside the scope of the current study, we can speculate on why SHS participants responded to and produced shorter word durations in their dominant L2 compared to L1-dominant speakers. It could be the case that heritage speakers in general are faster (faster RTs, shorter word durations) in their dominant language compared to L1-dominant speakers. To answer this question, one would need to measure a variety of L2 dominant English speakers' word productions of these stimuli and compare them to English controls. It could also be the case that these heritage speakers were simply faster to respond to stimuli compared to L1-dominant English participants due to social pressure. That is, culturally, heritage speakers are perceived as being in limbo with respect to a "native" language (Montrul, 2010a; 2012) despite the large body of data that shows they have native-like features in both languages. This speed-up observed in English could be an underlying compensatory method heritage speakers use to sound more native-like in their dominant L2. It does not seem to be the case that there is a global speed-up for these SHS as their word durations in Spanish are significantly lower compared to both their English and L1-dominant Spanish controls. The speed-up in English could also be a socio-phonetic marker of SHS English. SHS English has been shown to have a variety of unique acoustic features that distinguish this talker group from other varieties of English (e.g., Konopka, 2011). This increase in speed could then be the result of some kind of SHS English marker to indicate identity in that particular group. Finally, it could also be the case that this shortened duration is the result of a planning-production trade off. In some instances, we observed SHS having longer RTs compared to L1-dominant English controls (although this difference was not a significant main effect). As such, the shortened durations produced by SHS in English could be the result of more time being spent on word retrieval, resulting in less overall time for articulation. This trade-off would be consistent with Hennessey and Kirsner (1999) who have argued that pre-articulation processes (e.g., word planning) can modulate word duration, resulting in fewer resources (i.e., time) available for the production process.

Another area for future exploration is a qualitative analysis of the types of errors that SHS make in their non-dominant L1 Spanish. Recall that in general SHS were less intelligible and more adversely affected by some linguistic stresses in Spanish compared to L1-dominant Spanish controls. Additionally, heritage speakers were overall less accurate in correctly identifying real words in their non-dominant L1 Spanish compared to L1-dominant Spanish listeners. They were also less accurate in correctly rejecting non-words in both their dominant L2 English and their non-dominant L1 Spanish compared to L1-dominant controls in both languages. Where do these errors occur and do they fall into some kind of natural syntactic or semantic class of words? For example, in the speech intelligibility experiment (Experiment 1), if there were any consistent patterns (such as part of speech, open versus closed class, etc.) of words that consistently were marked as unintelligible for SHS, it would add to the subtle limitations of early exposure in language acquisition. With respect to experiment 2 (lexical decision), analyzing which words were consistently marked as false negatives (i.e., real words of Spanish that were identified by SHS as non-words) could provide a clearer picture of heritage speakers' language experience. While at ceiling-level accuracy in English real-word identification, SHS were significantly worse at accurately defining real words of Spanish compared to L1-dominant Spanish controls. These Spanish errors centered on low frequency and late-acquired words (see Table 6) with most errors in the late-acquired, low frequency cell. These data mirror, at least descriptively, mirror the reaction time results in which SHS received a benefit (i.e., faster RT) of early AoA or high frequency, but were more adversely affected by late-acquired or low-frequency forms in the non-dominant L1. While the current study lacks the statistical power to fully investigate this difference, such an analysis would provide unique insight into the successes and failures of language learning more generally. The lack of real-word identification errors in English (SHS and L1-dominant English controls achieved near-ceiling accuracy on real-word identification in English) further reinforces the idea that extended use is a means of overcoming any loss incurred by delayed acquisition to the dominant language.

Finally, as mentioned before, late second language (L2) learners should be included in future studies to show the qualitative differences between a non-dominant L1 and a non-dominant L2 speech learning process. While a large body of work already exists (Montrul and Foote, 204; Jegerski, 2016; Zamora, 2017, etc.) comparing the two, few studies have examined more global aspects of speech learning (sentence and word level production and perception) comparing late L2 learners and heritage speakers. Another comparison that exists between heritage speakers and late L2 learners is that in both cases, the target language (here the non-dominant L1 for HS and the non-dominant L2 for late learners) is subject to reduced variability in the input. For heritage speakers, this reduced variability stems from only a handful of speakers at home actively using the heritage language; for late L2 learners, reduced variability is due to classroom settings with only a small amount of instructors for each student. While talker and (other) acoustic variability in monolingual (L1-dominant) word identification has been shown to be either an impairment (e.g., Bent and Frush Holt, 2013) or a variable influence (e.g., Bradlow, Nygaard, and Pisoni, 1999; Sommers and Barcroft, 2006), it has actually been shown to be beneficial to successful second language acquisition (e.g., Barcroft 2004, 2007; Barcroft and Sommers, 2005), yet neither heritage speakers nor L2 learners actually receive high talker and/or acoustic variability in their non-dominant language. By comparing HS to late L2 learners, we may be able to get a sense of the influence of limited talker and/or acoustic variability in a direct comparison of (non-dominant) first and second language acquisition. That is, the reduced variability that heritage speakers received actually provided a means of achieving some L1-

dominant levels of performance in their L1 due to language learning mechanisms preferring a simpler, but early, input. At later stages of language learning when early input is no longer a viable option, extended, variable input becomes important. This result would then not only account for the results in Barcroft and Sommers (2005) showing increased L2 accuracy with high variability input, but also the results from simulations (discussed above) that have shown that reduced input is essential for first/early language acquisition (e.g., Elman, 1993). It could then explain how switched-dominance bilingualism successfully manifests in heritage speakers. The high variability that comes from extended use throughout adolescence and adulthood of their second language was actually key in achieving switched-dominance bilingualism. As such, the addition and comparison of late L2 learners to heritage speakers would also add to the existing research on the quality and timing of linguistic input on language acquisition.

5.7 Conclusion

The goal of the current study was to determine how language learning mechanisms operate on two levels of linguistic representation that have been shown to be sensitive to the order (i.e., age of acquisition) and amount (i.e., extended use resulting in language dominance) of language exposure, namely speech production and vocabulary retrieval and production. The switched-dominance bilinguals, referred to as heritage speakers, in this study provided unique insight into the subtle differences in language processing not readily available in monolingual or bilingual speakers who confound their early-acquired language (L1) with their dominant language. Results showed that heritage speakers were able to benefit from early exposure in conditions that supported successful linguistic communication (low levels of noise, early-acquired and high frequency words), highlighting that language learning mechanisms indeed rely on very little input early on to achieve native-like success. However, this success was limited in their non-dominant L1 Spanish, when the linguistic system was stressed. As such, in cases where their non-dominant L1 speech was placed in relatively higher levels of background noise (experiment 1) or when words in the non-dominant L1 were late-acquired or low frequency (experiments 2 and 3), heritage speakers failed to perform at the same level as an L1-dominant Spanish control. Conversely, heritage speakers had little to no difficulty in speech production and/or lexical processing in their dominant L2 (second-acquired) English compared to L1 dominant English participants even in the aforementioned more difficult language processing conditions. These results suggest that while language learning mechanisms require little input in the early stages of acquisition, extended use later on can also achieve the same native-like result and in fact, may be necessary for successful language acquisition to be maintained throughout adulthood. Combined, this study has enhanced the definition of what it means to be a “native” speaker of a language, such that focusing on an L1 versus L2 distinction, without any focus on dominant language status, may actually mislead and incorrectly predict how a particular speaker will behave under a given set of linguistic conditions.

References

Albirini, A., Benmamoun, E., & Saadah, E. (2011). Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Studies in Second Language Acquisition*, 33(2), 273-303.

Alonso, M. A., Fernandez, A., & Díez, E. (2014). Subjective age-of-acquisition norms for 7,039 Spanish words. *Behavior research methods*, 1-7.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32(2), 233-250.

Altmann, G. T. M., & Dienes, Z. (1999). Rule learning by seven-month-old infants and neural networks. *Science*, 284, 875.

Au, T. K. F., Knightly, L. M., Jun, S. A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological science*, 13(3), 238-243.

Baker, R. E., & Bradlow, A. R. (2007). Second mention reduction in Indian, English, and Korean. *Journal of the Acoustical Society of America*, 122(5), 2993.

Barcroft, J. (2004). Second language vocabulary acquisition: A lexical input processing approach. *Foreign Language Annals*, *37*(2), 200-208.

Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning*, *57*(1), 35-56.

Barcroft, J., & Sommers, M. S. (2005). Effects of acoustic variability on second language vocabulary learning. *Studies in Second Language Acquisition*,*27*(03), 387-414.

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2012). Random effects structures in mixed-effects models: Keep it maximal. *Journal of Memory and Language*.

Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: evidence for a lexical-semantic competition hypothesis. *Cognition*, 96(2), B45-B54.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92-111.

Bent, T., & Frush Holt, R. (2013). The influence of talker and foreign-accent variability on spoken word identification. *The Journal of the Acoustical Society of America*, 133(3), 1677-1686.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words*, *167*, 224.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*,*109*(2), 775-794.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 210-223.

Blake, R. J., & Zyzik, E. C. (2003). Who's Helping Whom?: Learner/Heritage‐Speakers' Networked Discussions in Spanish. *Applied Linguistics*, *24*(4), 519-544.

Blasingame, M., & Bradlow, A.R., (in prep). "Speech learning by switched-dominance bilinguals: Second language versus non-dominant language deficits"

Boersma, Paul & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.31, retrieved 21 August 2017 from http://www.praat.org/

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2001). Age-of acquisition and word frequency in the lexical decision task: Further evidence from the French language. *Cahiers de Psychologie Cognitive - Current Psychology of Cognition*, 20 (6), 401–443.

Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology*, 12(2), 189-221.

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29-49.

Bradlow, A. R., Ackerman, L., Burchfield, L. A., Hesterberg, L., Luque, J. S., & Mok, K. (2010). ALLSSTAR: Archive of L1 and L2 Scripted and Spontaneous Transcripts And Recordings. Department of Linguistics, Northwestern University.

Bradlow, A. R., Kim, M., & Blasingame, M. (2017). Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *The Journal of the Acoustical Society of America*, 141(2), 886-899.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Attention, Perception, & Psychophysics*, 61(2), 206-219.

Bradlow, A.R., Blasingame, M., & Lee, K. (in prep). Language-independent talker-specificity in bilingual speech intelligibility: Individual traits persist across L1 and L2 states. Submitted to *The Journal of Laboratory Phonology.*

Brysbaert, M. & New, B. (2009). Moving beyond Kucera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. *Behavior Research Methods*, 41 (4), 977-990.

Carlson, M. T., Goldrick, M., Blasingame, M., & Fink, A. (2016). Navigating conflicting phonotactic constraints in bilingual speech perception. *Bilingualism: Language and Cognition*, 19(5), 939-954.

Carroll, J. B., & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior*, 12(5), 563-576.

Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America*, 129(6), 3964-3980.

Chung, E. (2018). Second and heritage language acquisition of Korean case drop. *Bilingualism: Language and Cognition*, 21(1), 63-79.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *The Quarterly Journal of Experimental Psychology*, 60(8), 1072-1082.

Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 650-658.

Dirix, N., & Duyck, W. (2017a). An eye movement corpus study of the age-of-acquisition effect. *Psychonomic Bulletin & Review*, 1-7.

Dirix, N., & Duyck, W. (2017b). The first-and second-language age of acquisition effect in first-and second-language book reading. *Journal of Memory and Language*, 97, 103-120.

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.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition,* 48(1), 71-99.

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur-Bates communicative development inventories*. Paul H. Brookes Publishing Company.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474-496.

Gathercole, V. C. M. (2017). Straw man: Who thought exposure was the ONLY factor?. *Bilingualism: Language and Cognition*, 20(1), 23-24.

Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 267.

Gerhand, S., & Barry, C. (1999). Age-of-acquisition and frequency effects in speeded word naming. *Cognition*, 73(2), B27-B36.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. Acta *Psychologica*, 115(1), 43-67.

Godson, L. (2004). Vowel production in the speech of Western Armenian heritage speakers. *Heritage Language Journal*, *2*(1), 44-69.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1166.

Gollan, T. H., Ferreira, V. S., Cera, C., & Flett, S. (2013). Translation-priming effects on tip-of-the-tongue states. *Language and Cognitive Processes*, (1-22).

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of memory and language*, 58(3), 787-814.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General,* 140(2), 186.

Gharibi, K., & Boers, F. (2017). Influential factors in incomplete acquisition and attrition of young heritage speakers' vocabulary knowledge. *Language Acquisition*, 24(1), 52-69.

Haastrup, K., & Henriksen, B. (2001). The interrelationship between vocabulary acquisition theory and general SLA research. *EUROSLA yearbook,* 1(1), 69-78.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(04), 495-522.

Hennessey, N. W., & Kirsner, K. (1999). The role of sub-lexical orthography in naming: a performance and acoustic analysis. Acta psychologica, 103(1-2), 125-148.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production?. *Acta psychologica*, 127(2), 277-288.

Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica: Revista de metodología y psicología experimental*, 23(2), 245-282.

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language*, 64(1), 32-58.

Jegerski, J. (2016). The processing of the object marker a by heritage Spanish speakers. International Journal of Bilingualism, 1367006916681083.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. *Talker variability in speech processing*, 145-165.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*, *21*(1), 60-99.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42(3), 627-633.

Konopka, K. (2011). The Vowels of Mexican Heritage English: Language contact and vocalic variation in a Chicago community (Unpublished doctoral dissertation). Northwestern University. Evanston, USA.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44(4), 978-990.

Lewis, M. B., Gerhand, S., & Ellis, H. D., 2000. Re-evaluating age-of-acquisition effects: are they simply cumulative-frequency effects?. *Cognition*, 78(2), 189-205.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. PloS one, 7(8), e43230.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of speech, language, and hearing research,* 40(3), 686-693.

Montrul, S. (2010a). Current issues in heritage language acquisition. *Annual Review of Applied Linguistics*, 30, 3-23.

Montrul, S. (2010b). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research*, 26(3), 293-327.

Montrul, S. A. (2012). Is the heritage language like a second language?. *Eurosla Yearbook*, *12*(1), 1-29.

Montrul, S., Davidson, J., De La Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 17(1), 118-138.

Montrul, S., & Foote, R. (2014). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers*. International Journal of bilingualism*, 18(3), 274-303.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1), 116.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, 91(2), 167-180.

Morrison, C. M., Hirsh, K. W., Chappell, T., & Ellis, A. W. (2002). Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology*, 14(4), 435-459.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive* *psychology*, 47(2), 204-238.

Oh, J. S., Jun, S.-A., Knightly, L. M. & Au, T. K. (2003). Holding on to childhood language memory. *Cognition,* 86, B53–B64

Oh, S., & Lee, Y. (2016). Repeated Mention Reduction in L2 English Spontaneous Speech. *Speech Prosody* 2016, 961-964.

Pallier, C., Dehaene, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first?. *Cerebral cortex*, 13(2), 155-161.

Parlato-Oliveira, E., Christophe, A., Hirose, Y., & Dupoux, E. (2010). Plasticity of illusory vowel perception in Brazilian-Japanese bilinguals. *The Journal of the Acoustical Society of America*, 127(6), 3738-3748.

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current opinion in neurobiology*, 15(2), 202-206.

Pérez, M. A. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *The Quarterly Journal of Experimental Psychology*, 60(1), 32–42.

Peyton, J. K., Ranard, D. A., & McGinnis, S. (2001). *Heritage Languages in America: Preserving a National Resource*. Language in Education: Theory and Practice. Delta Systems Company Inc., 1400 Miller Parkway, McHenry, IL 60050-7030.

Pierrehumbert, J. B. (2001). Lenition and contrast. *Frequency and the emergence of linguistic structure*, 45, 137.

Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics*, *34*(4), 516-530.

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, 6(1), 40-71.

Polinsky, M., & Kagan, O. (2007). Heritage languages: In the 'wild'and in the classroom. *Language and Linguistics Compass*, *1*(5), 368-395.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Reid, J. M. (1995). *Learning styles in the ESL/EFL classroom*. Heinle & Heinle Publishers, International Thomson Publishing Book Distribution Center, 7625 Empire Drive, Florence, KY 41042.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, 27(03), 465-485.

Scherag, A., Demuth, L., Rösler, F., Neville, H. J., & Röder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition*, 93(3), B97-B108.

Soli, S. D., & Wong, L. L. (2008). Assessment of speech intelligibility in noise with the Hearing in Noise Test. *International Journal of Audiology*, 47(6), 356.

Sommers, M. S., & Barcroft, J. (2006). Stimulus variability and the phonetic relevance hypothesis: Effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification. *The Journal of the Acoustical Society of America*, *119*(4), 2406-2416.

Taft, M. (1986). Lexical access codes in visual and auditory word recognition. *Language and Cognitive Processes*, 1(4), 297-308.

Vlach, H. A. (2014). The spacing effect in children's generalization of knowledge: allowing children time to forget promotes their ability to learn. *Child Development Perspectives*, 8(3), 163-168.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental psychobiology*, 46(3), 233-251.

Zamora, C. C. (2017). The secret is in the processing: A study of levels of explicit computerized feedback in heritage and L2 learners of Spanish (Doctoral dissertation, Georgetown University).

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and language*, 47(1), 1-29.

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition*, 32(1), 31-38.

Appendix A: List of Experiment 1 Stimuli (HINT Sentences; Soli and Wong, 2008)

English Sentences:

A boy fell from the window.

The wife helped her husband.

Big dogs can be dangerous.

The shoes were very dirty.

The player lost a shoe.

Somebody stole the money.

The fire was very hot.

She's drinking from her own cup.

The picture came from a book.

The car is going too fast.

The paint dripped on the ground.

The towel fell on the floor.

The family likes fish.

The bananas are too ripe.

He grew lots of vegetables.

She argues with her sister.

The kitchen window was clean.

He hung up his raincoat.

The mailman brought a letter.

The mother heard the baby.

She found her purse in the trash.

The table has three legs.

The children waved at the train.

Her coat is on a chair.

The girl is fixing her dress.

It's time to go to bed.

Mother read the instructions.

The dog is eating some meat.

Father forgot the bread.

The road goes up a hill.

The painter uses a brush.

The family bought a house.

Swimmers can hold their breath.

She cut the steak with her knife.

They're pushing an old car.

The food is expensive.

The children are walking home.

They had two empty bottles.

Milk comes in a carton.

The dog sleeps in a basket.

The house had nine bedrooms.

They're playing in the park.

Rain is good for trees.

They sat on a wooden bench.

The child drank some fresh milk.

The baby slept all night.

The salt shaker is empty.

The policeman knows the way.

The buckets fill up quickly.

The boy is running away.

A towel is near the sink.

Flowers can grow in the pot.

He's skating with his friend.

The janitor swept the floor.

The lady washed the shirt.

She took off her fur coat.

The matchboxes are empty.

The man is painting a sign.

The dog came home at last.

They heard a funny noise.

They found his brother hiding.

The dog played with a stick.

The book tells a story.

The matches are on a shelf.

The milk was by the front door.

The broom was in the corner.

The new road is on the map.

She lost her credit card.

The boy did a handstand.

They took some food outside.

The young people are dancing.

They waited for an hour.

The shirts are in the closet.

They watched the scary movie.

The milk is in a pitcher.

The truck drove up the road.

The tall man tied his shoes.

A letter fell on the floor.

The ball bounced very high.

Mother cut the birthday cake.

The football game is over.

She stood near the window.

The kitchen clock was wrong.

The children helped their teacher.

They carried some shopping bags.

Someone is crossing the road.

She uses her spoon to eat.

The cat lay on the bed.

They're running past the house.

He's washing his face with soap.

The dog is chasing the cat.

The milkman drives a small truck.

The bus leaves before the train.

The baby has blue eyes.

The bag fell off the shelf.

They are coming for dinner.

They wanted some potatoes.

They knocked on the window.

School got out early today.

The football hit the goalpost.

The boy ran away from school.

Sugar is very sweet.

The two children are laughing.

The firetruck is coming.

Mother got a sauce pan.

The baby wants his bottle.

The ball broke the window.

There was a bad train wreck.

The waiter brought the cream.

The teapot is very hot.

The apple pie is good.

The jelly jar was full.

The girl is washing her hair.

The girl played with the baby.

The cow is milked every day.

They called an ambulance.

They are drinking coffee.

He climbed up the ladder.

*Spanish Sentences (*English Gloss in *Italics*)*:*

El niño hace ruido en su cuarto. (*The boy makes noise in his room*)

Él prefiere tomar el desayuno en el comedor. (*He prefers to have breakfast in the dining room*)

Yo recibí una carta hoy. *(I received a letter today*)

Él come salchichas con mostaza. (*He eats sausages with mustard*)

Él tomó café después de levantarse. (*He has coffee after getting up*)

Las noticias no son siempre buenas. (*News is not always good*)

Estuvimos esperando por dos horas. (*We waited for two hours*)

La cocina estaba llena de hormigas. (*The kitchen is full of ants*)

El avión despegó al amanecer. (*The airplane took off at dawn)*

Necesitas un pasaporte para volver al país. (*You need a passport to return to the country)*

Él pagó su cuenta en efectivo. (*He paid his bill in cash)*

La gallina saltó la cerca de el corral. (*The hen jumped over the corral fence)*

Su hermano se quedó a cenar. (*His brother stayed for dinner)*

El perro está comiendo carne. (*The dog is eating meat)*

El niño menor pateó la pelota. (*The younger boy hit the ball)*

La bailarina estaba muy cansada. (*The dancer was very tired)*

La joven recibió un collar de perlas. (*The young woman received a pearl necklace)*

Ella prometió regresar muy pronto. (*She promised to return very soon)*

Mi abuela me regaló un par de pantalones. (*My grandmother gifted me a pair of pants)*

Hoy hace mucho calor. (*It's very hot today)*

Su amiga está en el hospital. (*Her friend is in the hospital)*

Ella pasea por el parque. *(She goes around the park*)

Los niños juegan con el perro. (*The children play with the dog)*

Las montañas están cerca de la playa. *(The mountains are near the beach)*

El niño está tomando una limonada. *(The boy is having lemonade)*

Ella tiene mucho calor también. *(She is also very hot)*

La muchacha se cepilla los dientes. *(The girl brushes her teeth)*

El niño no quiere jugar hoy. (*The boy does not want to play today)*

El piso está cubierto de hojas. *(The floor is covered in leaves)*

La mamá gallina protegió sus huevos. *(The mother hen protected her eggs)*

La familia compró la casa. *(The family bought the house)*

A mí me gusta la sopa de verduras. *(I like vegetable soup)*

Ayer me caí de la bicicleta. *(Yesterday, I fell off my bicycle)*

Ellos trabajan en el estadio. *(They were working in the stadium)*

Los pasajeros están cansados de esperar. *(The passengers were tired of waiting)*

El pasajero olvidó su pasaporte. *(The passenger forgot his passport)*

El perro está ladrando muy fuerte. *(The dog is barking very loudly)*

La iglesia está cerca del mercado. *(The church is near the market)*

Ellos escucharon música en el parque. *(They listened to music at the park)*

Ese vestido verde cuesta mucho dinero. *(That green dress costs a lot of money)*

El piso está muy duro. *(The floor is very strong)*

Los hombres generalmente usan pantalones largos. *(Men usually wear long pants)*

El soldado estaba herido. *(The soldier was wounded)*

El policía lo conoció de inmediato. *(The police knew him immediately)*

Mi mamá trabaja con computadoras. *(My mother works with computers)*

Las naranjas también son frutas. *(Oranges are also fruits)*

Es posible que llueva hoy. *(It might rain today)*

El hombre se quitó su sombrero. *(The man took off his hat)*

La mujer está preparando verduras. *(The woman is preparing vegetables)*

Finalmente encontró a su hermano. *(He finally found his brother)*

El camión lleva fruta fresca. *(The truck brings fresh fruit)*

La señora está sentada en su silla. *(The woman is seated in her chair)*

Ellos invitaron a unos amigos a cenar. *(They invited some friends to dinner)*

El tren está viajando muy rápido. *(The train is traveling very fast)*

El papá olvidó sus llaves. *(The father forgot his keys)*

El niño se cayó de la escalera. *(The boy fell down the stairs)*

La princesa se casó con su sirvienta. *(The princess married her servant)*

Las orejas del ratón son enormes. *(Rats' ears are enormous)*

La niña compró helados. *(The girl bought ice cream)*

No me gusta cuando llueve. *(I don't like when it rains)*

La casa tiene un jardín hermoso. *(The house has a beautiful garden)*

El niño entró por la ventana. *(The boy entered through the window)*

El cocinero necesita más zanahorias. *(The cook needs more carrots)*

La mamá se quebró el tobillo derecho. *(The mother broke her right ankle)*

Las estrellas brillan de noche. *(The stars shine at night)*

Yo visito a mi abuela todos los días. *(I visit my grandmother every day)*

Ella se acordó de su amiga. *(She remembered her friend)*

A ella le gustan las novelas románticas. *(She likes romance novels)*

Había agua por todos lados. *(There was water on all sides)*

El niño juega con su gato. *(The boy plays with his cat)*

Los niños grandes salieron al patio. *(The big boys left the patio)*

Tomamos café en el desayuno. *(We have coffee at breakfast)*

La niña está buscando su muñeca. *(The girl is looking for her doll)*

Se le rompió el reloj esta mañana. *(The clock broke this morning)*

Los muchachos juegan al fútbol. *(The boys play soccer)*

Los toros son animales muy fuertes. *(Bulls are very strong animals)*

El padre se mira en el espejo. *(The father saw himself in the mirror)*

La pelota quebró el espejo. *(The ball broke the mirror)*

El mar está muy hondo. *(The sea is very deep)*

La fiesta está muy aburrida. *(The party is very boring)*

Ellos están comprando ropa para la escuela. *(They are buying school clothes)*

Yo me lavo los dientes antes de dormir. *(I brush my teeth before bed)*

Mamá nos lleva al colegio. *(Mother brings us to school)*

El sol brilla en el cielo. *(The sun shines in the sky)*

El equipo quiere ganar y no perder. *(The team wants to win and not lose)*

Los vaqueros cuidan al ganado. *(The cowboys look after the cattle)*

Los estudiantes visitaron el museo. *(The students visited the museum)*

La mamá compró frutas y verduras. *(The mother bought fruit and vegetables)*

Este edificio tiene cien metros de altura. *(This building measures 100 meters tall)*

A él no le gusta el pescado. *(He doesn't like fish)*

El agua del río está tibia. *(The river water is warm)*

El niño se cayó de el árbol. *(The boy fell from the tree)*

Ellos cantaron toda la noche. *(They sang all night)*

Él perdió su corbata ayer. *(He lost his tie yesterday)*

Su abrigo está en la silla. *(His coat is on the chair)*

El hombre está pintando el letrero. *(The man is painting the sign)*

La bruja trabajaba con nuevos hechizos. *(The witch was working with new spells)*

El policía ayudó a el señor. *(The police helped the man)*

Papá apagó sólo el fuego. *(Father put out the only fire)*

Mi hermano tiene muchos amigos interesantes. *(My brother has many interesting friends)*

Algunas víboras no son venenosas. *(Some snakes are not poisonous)*

El señor pintó la casa. *(The man painted the house)*

Mi abuela se compró una cartera. *(My grandmother bought a purse)*

Hace tres meses que no va al cine. *(It's been three months since he's gone to the movies)*

Él pagó con dinero americano. *(He paid with American money)*

Los deportes son muy populares. *(Sports are very popular)*

Su casa está en el bosque. *(His house is in the forest)*

El mantel amarillo cubre la mesa. *(The yellow tablecloth covers the table)*

El hospital está muy lejos. *(The hospital is very far)*

Su corazón latía muy rápido. *(His heart was beating rapidly)*

La señora trabaja en una oficina. *(The woman works in an office)*

El equipo estaba jugando bien. *(The team was playing well)*

La señora se puso un abrigo. *(The woman put on a coat)*

Las dos niñas se están riendo. *(The two girls are laughing)*

El pájaro voló sobre el mar. *(The bird flew over the sea)*

La niña estaba contenta. *(The girl was happy)*

Ella se quitó su abrigo de piel. *(She took off her leather coat)*

El castillo es viejo y espantoso. *(This castle is old and dreadful)*

Ella se fue con su hermano. *(She left with her brother)*

Ella es más alta que su hermano. *(She is taller than her brother)*

Appendix B: List of Experiment 2 and 3 Stimuli

Real English and Spanish words with age of acquisition (AoA: early/late; AoA in years), Frequency (Freq: high/low; LogF: log frequency), number of syllables (Syllables), Experiment status (A = Experiment 2 only, B = both experiments 2 and 3), orthographic length (Ortho_L) and phonological length (Phono_L)

English Stimuli

Word	AoA	Freq	Years	LogF	Syllables	Experiment	Ortho_L	Phono_L
abdomen	late	low	8.61	0.64	3	B	7	7
achievement	late	low	8.80	0.65	3	B	11	8
afternoon	early	high	4.65	1.97	3	A	9	7
agent	late	high	9.55	2.02	2	A	5	5
aggression	late	low	9.30	0.58	3	A	10	7
angel	early	high	4.00	1.90	2	A	5	4
attendant	late	low	9.79	0.61	3	A	9	8
avalanche	late	low	9.60	0.64	3	B	9	7
bib	early	low	4.53	0.17	1	B	3	3
biscuit	early	low	4.63	0.68	2	B	7	6
boat	early	high	3.84	1.99	1	A	4	3
bomb	late	high	8.00	1.74	1	B	4	3
box	early	high	4.30	1.96	1	A	3	4
cartoon	early	low	3.11	0.68	2	A	7	5
cell	late	high	10.00	1.74	1	B	4	3
chalk	early	low	4.47	0.66	1	A	5	3
class	early	high	4.95	2.07	1	B	5	4
claw	early	low	4.70	0.73	1	B	4	3
client	late	high	11.05	1.73	2	B	6	6
clutch	late	low	10.24	0.54	1	B	6	4
code	late	high	8.11	1.73	1	B	4	3
college	late	high	8.37	1.94	2	A	7	5
collision	late	low	9.05	0.57	3	A	9	7
computer	late	high	9.70	1.78	2	B	8	8
cone	early	low	4.67	0.59	1	A	4	3
court	late	high	8.39	2.01	1	A	5	3
crayon	early	low	3.20	0.15	1	B	6	5
department	late	high	9.84	1.81	3	B	10	9
doctor	early	high	4.60	2.42	2	B	6	5
donation	late	low	9.33	0.65	3	B	8	7
dresser	early	low	4.28	0.66	2	A	7	5
drink	early	high	3.47	2.40	1	B	5	5

elf	early	low	4.32	0.68	1	A	3	3
evidence	late	high	10.58	1.93	3	A	8	7
experience	late	high	8.95	1.79	4	B	10	11
exterior	late	low	10.53	0.53	4	B	8	9
fellow	late	high	8.89	1.78	2	B	6	4
game	early	high	4.26	2.37	1	B	4	3
general	late	high	8.05	2.07	3	B	7	6
glass	early	high	4.47	1.79	1	B	5	4
grape	early	low	3.94	0.70	1	A	5	4
hamster	early	low	4.37	0.50	2	A	7	6
hat	early	high	3.33	1.81	1	B	3	3
haven	late	high	9.13	2.57	2	B	5	5
insect	early	low	4.75	0.62	2	A	6	6
judge	late	high	8.85	1.91	1	A	5	3
key	early	high	3.58	1.94	1	A	3	2
kite	early	low	4.58	0.52	1	A	4	3
knob	early	low	4.68	0.54	1	A	4	3
lady	early	high	3.68	2.34	2	B	4	4
letter	early	high	4.74	1.92	2	A	6	4
lettuce	early	low	4.28	0.64	2	A	7	5
lieutenant	late	high	9.50	2.02	3	A	10	8
line	early	high	4.85	2.32	1	B	4	3
lollipop	early	low	3.89	0.44	3	B	8	7
mat	early	low	4.62	0.65	1	A	3	3
melon	early	low	4.21	0.72	2	B	5	5
mitten	early	low	4.72	0.16	2	B	6	5
movie	early	high	3.56	2.09	2	B	5	4
murder	late	high	8.79	2.05	2	B	6	4
music	early	high	3.81	2.18	2	B	5	6
noodle	early	low	3.44	0.59	2	A	6	4
nuisance	late	low	9.59	0.62	2	A	8	7
number	early	high	3.94	2.38	2	B	6	5
observer	late	low	9.70	0.67	3	B	8	6
pacifier	early	low	4.56	0.14	4	B	8	7
packet	late	low	8.94	0.55	2	A	6	5
pajama	early	low	4.27	0.36	3	B	6	6
paper	early	high	4.00	2.02	2	A	5	4
patriot	late	low	10.14	0.56	3	A	7	8
person	early	high	4.67	2.33	2	B	6	5
picture	early	high	4.05	2.14	2	B	7	5
pinkie	early	low	3.21	0.43	2	B	6	5
pleasure	late	high	8.21	1.91	2	A	8	5

popsicle	early	low	3.78	0.42	3	B	8	7
position	late	high	8.72	1.86	3	A	8	7
potty	early	low	2.28	0.43	2	B	5	4
pressure	late	high	8.95	1.73	2	B	8	5
professor	late	high	10.89	1.85	3	A	9	7
relationship	late	high	9.11	1.84	4	A	12	10
respect	late	high	8.50	1.86	2	A	7	7
revolver	late	low	9.26	0.58	3	A	8	7
ride	early	high	4.67	2.13	1	B	4	3
ring	early	high	4.53	1.97	1	A	4	3
road	early	high	4.55	2.05	1	B	4	3
rock	early	high	3.22	1.94	1	A	4	3
roster	late	low	8.78	0.50	2	B	6	5
rumble	late	low	8.33	0.57	2	A	6	5
sake	late	high	8.74	1.81	1	B	4	3
saliva	late	low	8.42	0.56	3	A	6	6
scene	late	high	8.17	1.88	1	A	5	3
security	late	high	8.06	1.98	4	A	8	9
sergeant	late	high	9.00	1.81	2	B	8	6
service	late	high	8.00	1.91	2	A	7	5
sex	late	high	9.79	2.19	1	B	3	4
significance	late	low	10.11	0.65	4	B	12	11
simulation	late	low	10.53	0.51	4	B	10	10
situation	late	high	9.26	1.92	4	A	9	9
slate	late	low	9.25	0.63	1	A	5	4
spinach	early	low	4.94	0.55	2	A	7	6
splash	early	low	3.67	0.72	1	B	6	5
splinter	early	low	4.72	0.39	2	B	8	7
stallion	late	low	9.42	0.62	2	A	8	6
star	early	high	3.89	1.92	1	A	4	3
sticker	early	low	4.42	0.58	2	A	7	5
store	early	high	4.76	1.92	1	A	5	3
story	early	high	3.89	2.35	2	B	5	5
stretcher	late	low	9.39	0.51	2	B	9	6
stripe	early	low	4.05	0.39	1	B	6	5
summit	late	low	10.11	0.64	2	A	6	5
support	late	high	8.53	1.71	2	B	7	5
suspense	late	low	9.53	0.51	2	B	8	7
symphony	late	low	8.33	0.66	3	B	8	7
system	late	high	9.90	1.97	2	A	6	6
table	early	high	4.39	2.03	2	A	5	4
tape	early	high	4.42	1.84	1	A	4	3

terrain	late	low	10.95	0.51	2	B	7	5
torment	late	low	8.61	0.57	2	A	7	6
train	early	high	4.00	1.98	1	A	5	4
trench	late	low	10.21	0.61	1	A	6	5
twilight	late	low	10.00	0.61	1	A	8	6
vanilla	early	low	4.80	0.68	3	A	7	6
veil	late	low	9.69	0.60	1	A	4	3
veteran	late	low	9.90	0.68	3	B	7	7
voice	early	high	4.83	1.94	1	A	5	3
wall	early	high	3.79	1.86	1	A	4	3
womb	late	low	9.48	0.65	1	B	4	3
yogurt	early	low	3.72	0.13	2	B	6	5

Spanish Stimuli

Word	Gloss	AoA	Freq	Years	LogF	Syllables	Experiment	Ortho_L	Phono_L
abogado	lawyer	late	high	8.44	2.42	4	A	7	7
aborrecimiento	hatred	late	low	9.96	0.301	6	B	14	13
adobado	marinade	late	low	9.14	0.477	4	B	7	7
agudeza	intensity	late	low	9.78	0.301	4	B	7	7
alacena	cupboard	late	low	9.10	0.301	4	A	7	7
alcalde	mayor	late	high	8.02	2.272	3	B	7	7
amante	lover	late	high	8.96	1.748	3	B	6	6
anguila	eel	late	low	8.68	0.301	3	A	7	6
anhelo	desire	late	low	9.70	0.602	3	A	6	5
aparador	hutch	late	low	9.78	0.477	4	A	8	8
apio	celery	late	low	8.74	0.301	2	A	4	4
asunto	subject	late	high	8.16	2.303	3	A	6	6
avión	airplane	early	high	3.62	2.265	2	A	5	5
azufre	sulfur	late	low	9.44	0.477	3	A	6	6
basura	trash	early	high	3.56	1.806	2	A	6	6
bicho	bug	early	low	2.96	0.672	2	B	5	4
borrasca	low air	late	low	8.35	0.477	2	B	8	7
bravura	ferocity	late	low	9.78	0.301	3	B	7	7
brazo	arm	early	high	3.02	1.881	2	A	5	5
buhardilla	loft	late	low	9.40	0.602	4	B	10	8
calle	street	early	high	4.04	3.068	2	B	5	4
calor	heat	early	high	3.92	2.369	2	B	5	5
cama	bed	early	high	2.42	2.354	2	B	4	4
campaña	campaign	late	high	8.34	2.427	3	B	7	7
canción	song	early	high	3.40	2.461	2	B	7	7

cargo	position	late	high	8.74	2.137	2	A	5	5
carne	meat	early	high	4.16	2.045	2	A	5	5
carrito	cart	early	low	3.62	0.699	3	B	7	7
ceja	eyebrow	early	low	4.26	0.577	2	A	4	4
cepillo	brush	early	low	3.66	0.699	3	B	7	6
cielo	sky	early	high	3.58	2.093	2	A	5	5
cita	appointment	late	high	8.44	2.049	2	A	4	4
cobertizo	hut	late	low	9.00	0.602	4	B	9	9
cochecito	stroller	early	low	3.38	0.699	4	B	9	8
compañero	partner	early	high	4.66	2.362	4	B	9	9
compromiso	agreement	late	high	8.96	2.127	4	B	10	10
confianza	trust	late	high	8.3	2.272	3	A	9	9
consejo	advice	late	high	8.22	2.504	3	A	7	7
contorno	edge	late	low	8.38	0.477	3	B	8	8
corazón	heart	early	high	4.54	2.489	3	B	7	7
cuidado	caution	early	high	4.54	2.516	3	B	7	7
daño	damage	early	high	4.42	2.107	2	A	4	4
datos	facts	late	high	8.08	2.68	2	B	5	5
depilación	hair removal	late	low	9.24	0.477	4	A	10	10
discurso	speech	late	high	8.5	1.996	3	B	8	8
empleo	job/work	late	high	8.22	2.223	3	A	6	6
empresa	company	late	high	8.96	2.589	3	B	7	7
entereza	strength	late	low	9.94	0.301	4	B	8	8
época	era	late	high	8.02	2.734	3	B	5	5
escalón	step/stair	early	low	3.54	0.699	3	B	7	7
espalda	back	early	high	3.9	1.732	3	A	7	7
estado	state	late	high	9.72	3.01	3	B	6	6
estilo	style	late	high	8.76	2.26	3	A	6	6
estrella	star	early	high	3.66	2.104	3	A	8	7
estribo	stirrup	late	low	9.76	0.602	3	B	7	7
flaqueza	weakness	late	low	8.86	0.301	3	B	8	7
foca	seal	early	low	4.08	0.301	3	A	4	4
fresa	strawberry	early	low	3.18	0.699	2	B	5	5
fuego	fire	early	high	3.68	2.158	2	A	5	5
gobierno	government	late	high	8.88	2.999	3	A	8	8
guardería	nursery	early	low	3.46	0.680	4	B	9	9
gusano	worm	early	low	3.68	0.602	3	B	6	6
hada	fairy	early	low	3.86	0.677	2	B	4	3
hielo	ice	early	high	4.48	2.000	2	A	5	4
hormiga	ant	early	low	4.02	0.301	3	A	7	6
impresora	printer	late	low	9.24	0.602	2	A	9	9
informe	report	late	high	9.38	2.31	3	A	7	7

jabón	soap	early	low	3.16	0.570	2	B	5	5
jaqueca	migraine	late	low	8.90	0.301	3	A	7	6
jarabe	syrup	early	low	8.24	0.602	3	A	6	6
juez	judge	late	high	8.44	2.459	1	A	4	4
juicio	trial	late	high	8.7	2.438	2	A	6	6
jurado	jury	late	high	8.78	2.207	3	B	6	6
lágrima	teardrop	early	low	3.86	0.667	3	A	7	7
lavabo	sink	early	low	3.72	0.643	3	A	6	6
lente	lense	late	low	8.24	0.301	2	A	5	5
ley	law	late	high	8.54	2.919	1	A	3	2
lienzo	canvas	late	low	8.62	0.301	2	A	6	6
luz	light	early	high	3.18	2.534	1	B	3	3
maña	knack	late	low	8.84	0.477	2	B	4	4
mañana	morning	early	high	3.7	2.998	3	B	6	6
marcha	protest	late	high	8.52	2.461	2	B	6	5
mentira	lie	early	high	4.50	2.225	3	A	7	7
merienda	snack	early	low	3.74	0.409	3	A	8	8
miedo	fear	early	high	3.6	2.784	2	B	5	5
mimbre	wicker	late	low	8.78	0.602	2	A	6	6
mitad	half	early	high	4.96	2.542	2	B	5	5
moco	snot	early	low	2.84	0.477	2	A	4	4
moho	mildew	late	low	8.18	0.477	2	A	4	3
mugre	grime	late	low	8.82	0.301	2	A	5	5
navío	vessel	late	low	9.20	0.602	3	B	5	5
negocio	business	late	high	8.94	2.072	3	A	7	7
noche	night	early	high	3.22	3.001	2	B	5	4
ombligo	navel	early	low	4.62	0.477	3	A	7	7
pañal	diaper	early	low	4.24	0.675	2	B	5	5
pantorrilla	calf	late	low	8.32	0.477	4	A	11	10
papelera	garbage	early	low	3.88	0.397	4	A	8	8
paraguas	umbrella	early	low	3.56	0.677	3	B	8	8
payo	peasant	late	low	9.78	0.301	2	B	4	4
peine	comb	early	low	3.22	0.430	2	A	5	4
pelo	hair	early	high	2.60	2.286	2	A	4	4
peluche	plush toy	early	low	2.14	0.316	3	A	7	6
piojo	louse/lice	early	low	4.58	0.301	2	B	5	5
piragua	canoe	late	low	8.55	0.301	3	B	7	7
pregunta	question	early	high	4.6	2.94	3	B	8	8
pretendiente	candidate	late	low	8.74	0.477	4	B	12	12
propósito	intention	late	high	9.36	1.973	4	A	9	9
regalo	gift	early	high	2.82	2.124	3	A	6	6
reloj	clock	early	high	4.38	2.000	2	A	5	5

riesgo	risk	late	high	8.5	2.265	2	B	6	6
rizos	curls	early	low	4.78	0.477	2	A	5	5
salida	exit	early	high	4.96	2.505	3	B	6	6
servilleta	napkin	early	low	3.56	0.602	4	A	10	9
sonajero	rattle	early	low	3.24	0.699	4	B	8	8
sueño	dream	early	high	3.90	2.336	2	B	5	5
temporada	season	late	high	8.32	2.423	4	B	9	9
tenedor	fork	early	low	3.28	0.678	3	B	7	7
teniente	deputy	late	high	8.26	1.813	3	A	8	8
testigo	witness	late	high	8.56	1.886	3	B	7	7
tierra	earth	early	high	3.2	2.683	2	B	6	6
tijeras	scissors	early	low	3.66	0.602	3	B	7	7
tomo	collection	late	high	8.46	1.833	2	B	4	4
uña	fingernail	early	low	3.80	0.424	2	A	3	3
ventana	window	early	high	3.20	1.886	3	A	7	7
vestido	dress	early	high	3.08	1.978	3	A	7	7
viga	beam	late	low	8.80	0.477	2	A	4	4
voluntad	will (n.)	late	high	8.48	2.305	3	B	8	8
zanahoria	carrot	early	low	3.66	0.505	4	A	9	8

Appendix C: Experiment 2 Non-words With Input Word from Wuggy (Keuleers and Brysbaert, 2010)

Source word (Source), generated stimuli (non-word) and orthographic length (Ortho_L)

English Non-word Stimuli

Source	Non-word	Ortho_L
class	clade	5
doctor	puctor	6
drink	clish	5
game	gace	4
glass	wress	5
hat	vep	3
lady	tagy	4
line	lunk	4
line	lums	4
movie	soomie	6
music	mumic	5
number	nesser	6
person	purbon	6
picture	puffine	7
ride	samp	4
road	goom	4
story	stosy	5
bib	horm	4
biscuit	bisker	6
claw	clag	4
crayon	brigon	6
lollipop	mollipar	8
melon	muson	5
mitten	rallen	6
pacifier	madibier	8
jam	jahser	6
pinkie	pintie	6
pop	nuxel	5
potty	mattip	6
splash	throst	6
splinter	shrinter	8
stripe	strime	6
yoghurt	doskurt	7
bomb	boir	4

cell	vudge	5
client	spiemy	6
code	coss	4
computer	rommuter	8
department	depatemen	9
experience	enveriem	8
fellow	tullar	6
general	naneral	7
haven	hasem	5
murder	murper	6
pressure	chaggone	8
sake	rint	4
sergeant	serbeesh	8
sex	sef	3
support	suppome	7
abdomen	apnomer	7
achievement	acleavemic	10
avalanche	bavalaffe	9
clutch	brotch	6
donation	dapaterg	8
exterior	extenemior	10
observer	ersellent	9
roster	sestel	6
significant	pippificant	11
simulation	pimucation	10
stretcher	bratchen	8
suspense	nusperve	8
symphony	synclony	8
terrain	lerrake	7
womb	wect	4

Spanish Non-word Stimuli

Source	Non-word	Ortho_L
adobado	amopado	7
agudeza	abudeta	7
alcalde	arcalse	7
amante	anente	6
apio	abia	4
bicho	guirro	6
borrasca	borrinsa	8

bravura	cravuba	7
buhardilla	buordilla	10
calle	cache	5
calor	caler	5
cama	maba	4
campaña	tampiba	7
canción	candión	7
carrito	marrito	7
cepillo	cedallo	7
cobertizo	cofartivo	9
compañero	cospañero	9
compromiso	complorino	10
contorno	contonso	8
corazon	coramón	7
cuidado	caidado	7
discurso	sospirso	8
empresa	emtresa	7
entereza	enteneza	8
epoca	égoco	5
escalon	espalón	7
estado	escamo	6
estribo	estrabo	7
flaqueza	claquera	8
fresa	brena	5
gobierno	fopuerno	8
guardería	gonsmería	9
gusano	fusino	6
hada	gaca	4
jabón	fagón	5
jurado	jurido	6
jurado	zucido	6
jurado	zurado	6
lienzo	luespo	6
luz	huz	3
maña	mafa	4
mañana	pafina	6
marcha	parcho	6
miedo	miaco	5
mitad	mital	5
navio	napao	5
noche	golle	5
pañal	majal	5

paraguas	ragad	5
piragua	piramua	7
pregunta	preponta	8
pretendiente	pretergiante	12
riesgo	ruisgo	6
salida	ralada	6
sonajero	momabero	8
sueño	ruezo	5
temporada	bomporama	9
tenedor	tesedón	7
testigo	testizo	7
tierra	saulla	6
tijeras	biferas	7
tomo	sogo	4
voluntad	moluntar	8

Appendix D: Additional Figures from Experiment 2

Figure D1. Overall log reaction time (RT) for English controls and Spanish heritage speakers (SHS) in English.
Error bars represent 95% confidence intervals around the mean. There are no significant differences between English controls and SHS overall RT.

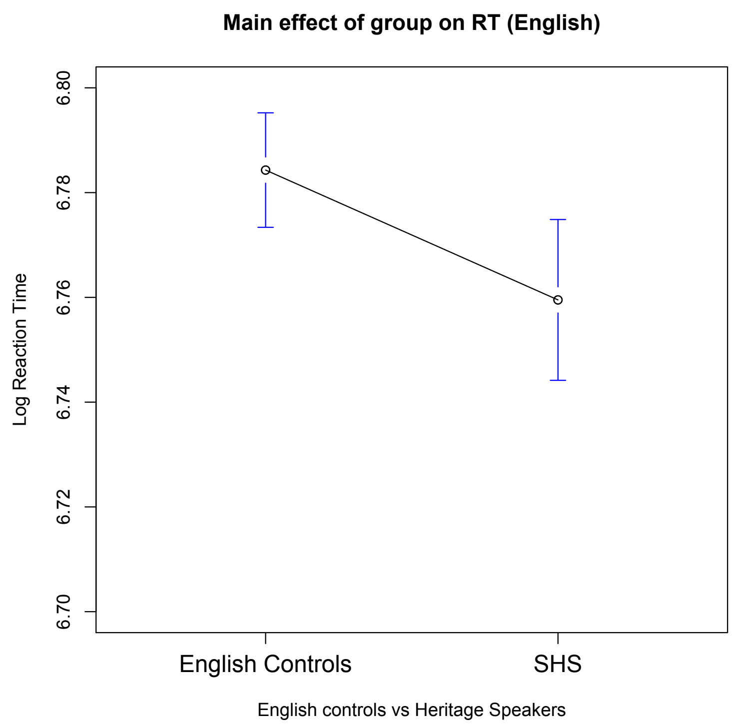


Figure D2. Interaction between Lexical Frequency and Group on English Word Reaction Time Interaction between lexical frequency (high versus low) and group (English control versus SHS) on log reaction time (RT) for English controls and Spanish heritage speakers (SHS) in English. Error bars represent 95% confidence intervals around the mean. SHS participants responded as fast as English controls to both high and low frequency words (no significant interaction).

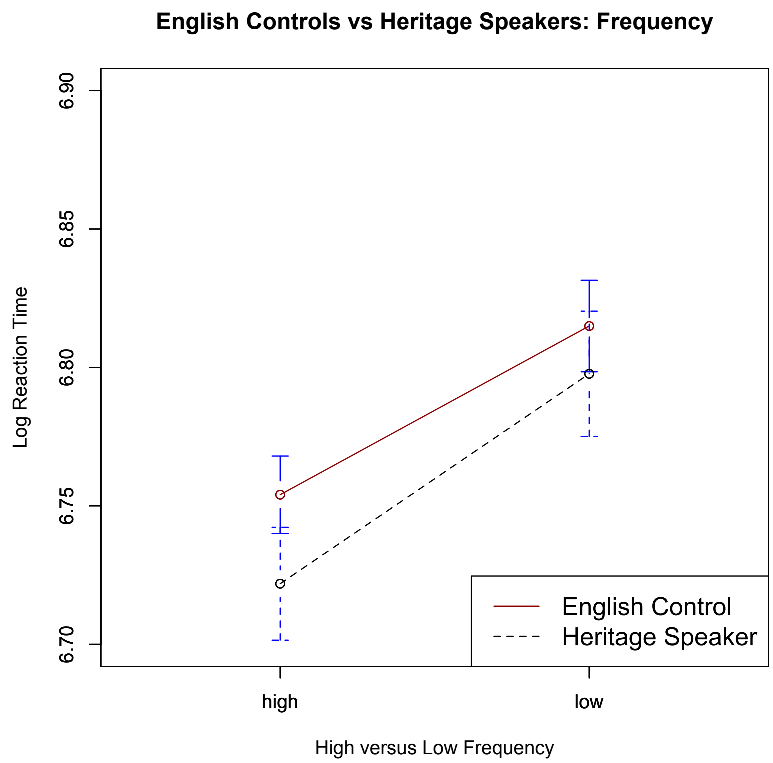


Figure D3. Log Reaction time (RT) for Spanish Controls and Spanish Heritage Speakers (SHS) in Spanish

Error bars represent 95% confidence intervals around the mean. There are no significant differences between Spanish controls and SHS overall RT. While this difference may appear significant, it is the interactions between group (SHS versus L1-dominant Spanish) and age of acquisition and lexical frequency (respectively) that drives this overall difference.

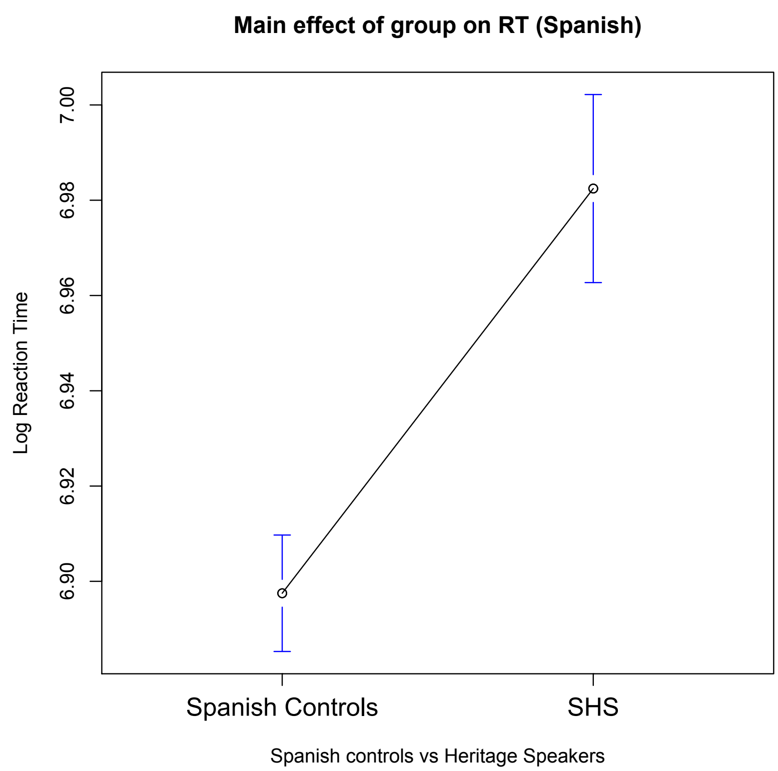
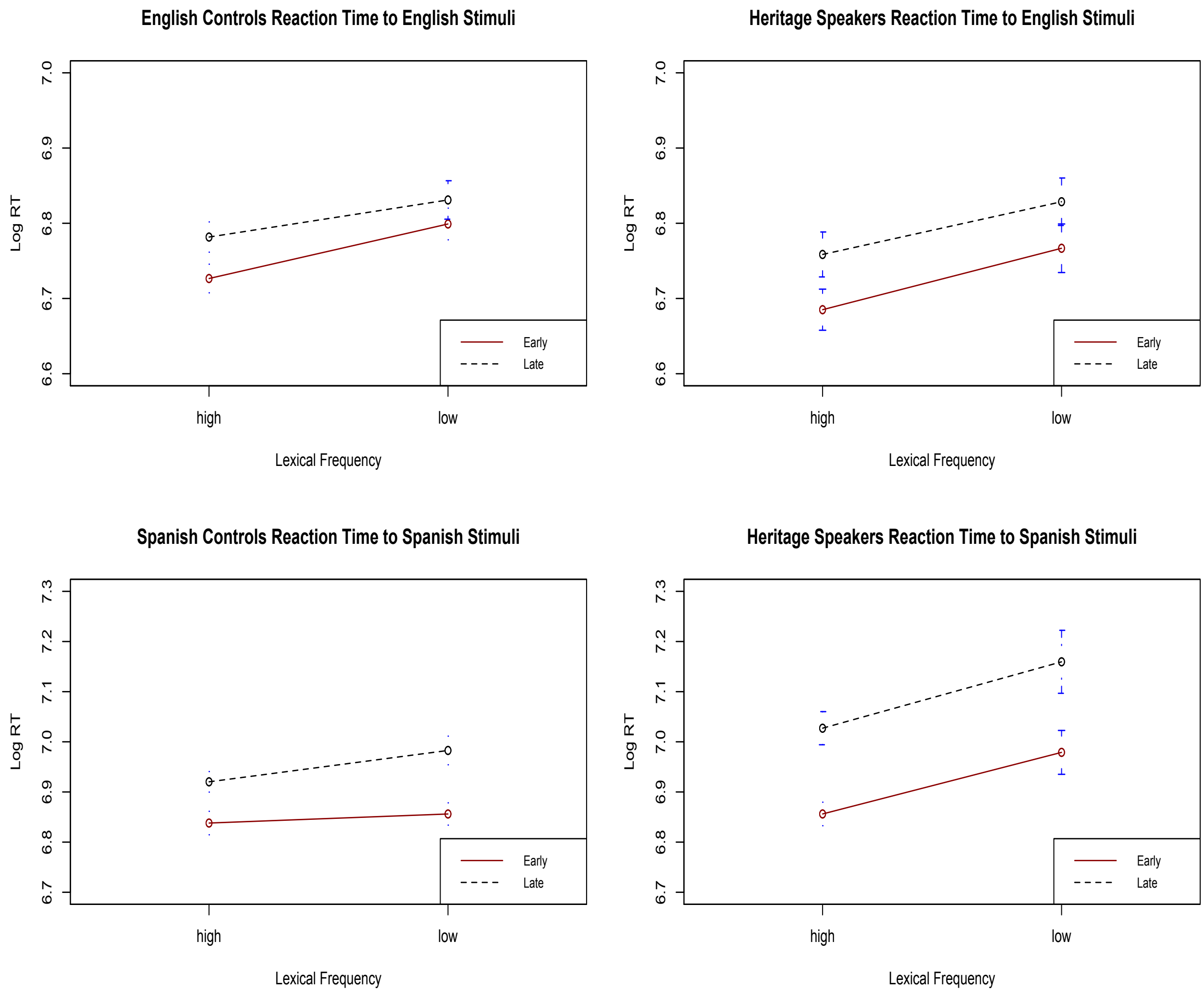


Figure D4. Individual Group Reaction Time Divided by Age of Acquisition and Lexical Frequency
This figure displays individual group reaction times (log RT) with respect to lexical frequency (high or low) and age of acquisition (AoA, early or late). Error bars represent 95% confidence intervals around the mean.



VITA

Michael Foxworth Blasingame
Department of Linguistics, Northwestern University
2016 Sheridan Road, Evanston, IL 60208 USA
Email: mblasingame [at] u [dot] northwestern [dot] edu

Education

2013-present: **Ph.D.** candidate in Linguistics, Northwestern University, Evanston, IL

- Cognitive science specialist
- *PhD Title: Early versus Extended Exposure in Speech and Vocabulary Learning: Evidence from Switched-dominance Bilinguals*
- Committee: Ann Bradlow, Matthew Goldrick, and Viorica Marian
- Date of completion: March 2018

2010-2013: **Master of the Arts** in Linguistics, Northwestern University, Evanston, IL

- *Master's Title: Speech-in-noise recognition by switched-dominance bilinguals: Second language versus non-dominant language deficits.*

2006-2010: **Bachelor of the Arts** in Linguistics and Spanish, University of Illinois at Urbana-Champaign

- *BA Thesis Title: Prosodic Effects on the Northern Cities Chain Shift in Chicago.*
- Translation (Spanish to English) certification program
- Chancellor's Scholar

Professional Skills

Programming languages: R, HTML, MaxMSP, Praat, SuperLab, XML
Natural languages: English (dominant), Spanish (advanced), Japanese (intermediate), Brazilian Portuguese (Intermediate), Sicilian (heritage)

Publications [in reverse chronological order]

Bradlow, A. R., Kim, M., & **Blasingame, M**. (2017). Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *The Journal of the Acoustical Society of America*, 141(2), 886-899.

Matsubara, J., **Blasingame, M.,** & Smith, E. A. (2017). An empirical investigation of the felicity conditions for the Japanese evidentials -rashii, -sooda, and -yooda. *International Review of Pragmatics*, 9(1)

Carlson, M. T., Goldrick, M., **Blasingame, M**., & Fink, A. (2016). Navigating conflicting phonotactic constraints in bilingual speech perception. *Bilingualism: Language and Cognition*, 19(5), 939-954.

Cole, J., Hualde, J.I., **Blasingame, M.,** Mo, Y. (2010). Shifting Chicago vowels: prosody and sound change. In Proceedings of Speech Prosody 2010, Chicago, IL

Presentations and Posters [in reverse chronological order]

Blasingame, M. and Bradlow. A. R. (2016). Early versus frequent exposure in speech production: evidence from heritage speakers. Annual meeting of the Linguistics Society of America, Washington, DC.

Blasingame, M., Bradlow, A.R. (2014). Switched-dominance bilingual speech production: Continuous usage versus early exposure. Accepted as a poster at the 168th Meeting of the Acoustic Society of America (Fall 2014) Indianapolis, Indiana.

Carlson, M., **Blasingame, M.,** Fink, A., Goldrick, M. (2014). Gradient phonotactic grammars in bilingual speech perception. Accepted at the 88th Annual Meeting of the Linguistic Society of America (2014) Minneapolis, Minnesota [34% acceptance rate]

Blasingame, M., Bradlow, A.R. (2013). Role of language dominance and early acquisition on speech learning in switched-dominance bilinguals. Accepted as a poster at the International Conference on Multilingualism (2013) McGill University, Montreal, Canada [57% acceptance rate]

Carlson, M., **Blasingame, M.,** Fink, A., Goldrick, M. (2013). Do you hear what I hear? Priming language-specific phonotactic constraints in speech perception. Accepted at the International Conference on Multilingualism (2013) McGill University, Montreal, Canada [57% acceptance rate]

Blasingame, M., Bradlow, A.R. (2013). First language versus dominant language intelligibility in "switched dominance" bilinguals. Presented at the Mid-Continental Phonetics and Phonology Conference (2013), Ann Arbor, MI

Matsubara, J., **Blasingame, M.,** Smith, E. A. (2012). Exploring felicitous environments for the Japanese evidential rashii . Accepted at the New Directions in Experimental Pragmatics workshop for the 13th conference of the International Pragmatics Association, New Dehli, India

Blasingame, M., Bradlow, A.R. (2012). Perception of speech-in-noise for L2 learners and heritage speakers in both L1 and L2 . Accepted at The 164th Meeting of the Acoustic Society of America (Fall 2012), Kansas City, USA

Luque, J., **Blasingame, M.,** Burchfield, L.A., Matsubara, J., Bradlow, A.R. (2012). The relationship between first language and second language intelligibility in Mandarin-English bilinguals. Accepted at The 164th Meeting of the Acoustic Society of America (Fall 2012), Kansas City, USA

Blasingame, M., Cole, J., Hualde, J.I., (2012). Competing effects of prominence and vowel shift in Chicago English. Accepted at The 13th Conference on Laboratory Phonology (2012), Stuttgart, Germany [50% acceptance rate]

Cooper, A. **Blasingame, M.,** Potter, D. (2011). Communicatively- and prosodically-driven hyper-articulation in English. Presented at the Mid-Continental Phonetics and Phonology Conference (2011), Champaign-Urbana, IL

Hualde, J.I., **Blasingame, M.,** Nadeu, M., and Simonet, M. (2009). "Effect of the preceding segment on Spanish spirantization." Presented at the Hispanic Linguistics Symposium in San Juan, Puerto Rico, October 2009

Research Experience and Grants [in reverse chronological order]

- Linguistics Department Research Grant (Fall 2015)
- Graduate Research Grant (Spring 2014)
- Research Assistantship (Winter, Spring 2013): National Science Foundation, Grant # BCS-0846147. "Integrating grammatical and psycholinguistic approaches to phonological processes in speech production."
- Research Assistantship (Fall 2012): National Institute for Deafness & Other Communication Disorders, Grant # R01 DC005794. "Talker-Listener Alignment during Speech Production and Perception."
- Research Assistant for Professor Jose Ignacio Hualde, PhD. (Fall 2009-Spring 2010)|
- Senior Supervisor for Professor Jennifer Cole, PhD. research project on prosody (Spring 2010)
- Undergraduate Research Assistant (Fall 2007 through Summer 2010) for Professor Kiel Christianson, PhD. (Department of Educational Psychology-Psycholinguistics)

Teaching

Instructor and guest lectures:

- International Summer Institute (ISI); English Conversation and Presentations; English Pronunciation; Summers: 2013-2017
- Instructor Linguistics 380: English Pronunciation (Fall 2014)
- Guest Lecture: Linguistics 450-2 (Laboratory Phonology 2/Acoustic Phonetics) Northwestern University (Spring 2014)
- Guest Lecture: Linguistics 321 (Bilingualism) Northwestern University (Spring 2014)
- Test Prep Instructor ISI (Summer 2012)
- Linguistics 380: English Conversation and Fluency (Winter 2012)

Individual tutor: English as a Second Language Program (Spring 2012 through Winter 2018)

Teaching Assistant:

- Teaching Assistant Linguistics 260: Introduction to Analysis of Words and Sentences (Winter 2017)
- Teaching Assistant Linguistics 250: Introduction to Sound Patterns of Human Language (Fall 2011, Winter 2016)

Service

- Graduate student liaison (2016-2017)
- Graduate Student Association Representative (2014-2016)
- Phonatics Coordinator (Fall 2013 through Summer 2014)
- Graduate Student Association Department Representative [Linguistics] (Fall 2012 through Summer 2014)
- Linguistics Noon Colloquium Coordinator (Fall 2013 through Summer 2014)
- Linguistics Department Social Chair (Fall 2011 through Summer 2013)
- Linguistics Noon Colloquium Coordinator (Fall 2011 through Summer 2012)
- Linguistics Colloquia Committee (Fall 2010 through Summer 2011)