THE PENNSYLVANIA STATE UNIVERSITY

The Graduate School

Department of Psychology

THE PLASTICITY OF THE NATIVE LANGUAGE IN ADULT SECOND LANGUAGE LEARNING

A Thesis in

Psychology

by

Kinsey Bice

© 2014 Kinsey Bice

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

December 2014
The thesis of Kinsey Bice was reviewed and approved* by the following:

Judith F. Kroll  
Distinguished Professor of Psychology, Linguistics, and Women's Studies  
Thesis Advisor

Paola E. Dussias  
Professor of Spanish, Linguistics and Psychology

Daniel J. Weiss  
Associate Professor of Psychology and Linguistics

Pamela M. Cole  
Liberal Arts Research Professor of Psychology, Human Development, and Family Studies  
Assistant Director of Clinical Training

Melvin M. Mark  
Professor of Psychology  
Head of the Department of Psychology

*Signatures are on file in the Graduate School
ABSTRACT

Learning a second language (L2) in adulthood is a difficult task that many attempt with varying levels of success. Past research has focused on hard maturational constraints as the constraining factor in adult L2 learning, but recent studies challenge those accounts by demonstrating that some adult L2 learners have achieved native-like processing in the L2. The present study shifts the focus in adult L2 learning from the L2 to the native language (L1), by testing a new hypothesis that the L1 must change in the process of L2 learning, and that the learners who are better able to tolerate L1 changes achieve higher L2 proficiency. Tolerance to L1 change incorporates two types of changes, L1 costs and L2 sensitivity. L1 costs are thought to be modulated by inhibitory control, whereby L2 learners who are better able to inhibit the L1 would have lower L1 performance compared to monolinguals, but overall higher L2 proficiency. The other prediction was the L2 learners who reveal the influence of the L2 during L1 processing, or L2 sensitivity, would have higher L2 proficiency. To test this hypothesis, L2 learners of various proficiency levels were tested in the L1 and the L2 on tasks of comprehension (lexical decision task) and production (semantic fluency task) and compared to monolinguals in the L1 tasks. ERP measurements were taken during the comprehension task to investigate neural sensitivity to the L2 that might not otherwise be present in behavioral measures. Statistical analyses revealed that L2 learners and monolinguals benefited from the presence of cognates in the L1, but the ERP waveforms and scalp distributions revealed that only the L2 learners demonstrated neural activity consistent with cognate facilitation, in the form of a reduced N400 for cognates compared to noncognates. Interestingly, the L2 results from the comprehension task showed that the beginning L2 learners did not produce an N400 for cognates compared to noncognates, but the advanced L2 learners had a clear N400 effect. Results are discussed in terms of the hypothesis and past research finding L2 sensitivity during L1 processing.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... vii

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

ACKNOWLEDGEMENTS ................................................................................................. xiii

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

Age of Acquisition and the Critical Period Hypothesis .............................................. 2
The Critical Period Hypothesis and Shallow Structure Hypothesis ............................. 2
Alternative Accounts to the Critical Period Hypothesis .............................................. 4
Native vs. Non-Native Language Processing ............................................................... 7
L1 Changes in Highly Proficient Bilinguals .............................................................. 11
Parallel Activation ....................................................................................................... 11
Influence of the L2 on the L1 ....................................................................................... 14
L1 Costs ......................................................................................................................... 18
The Present Study ........................................................................................................ 22
Overview of Tasks, Predictions, Planned Analyses ................................................... 27

Chapter 2 Methods ..................................................................................................... 31

Participants .................................................................................................................. 31
Materials ...................................................................................................................... 32
Procedure ..................................................................................................................... 35

Chapter 3 Results ..................................................................................................... 39

Data Cleaning and Coding Procedures ....................................................................... 39
Group Comparisons .................................................................................................... 42
English Lexical Decision Task .................................................................................. 42
Words vs. Nonwords ................................................................................................. 44
Cognates vs. Noncognates ....................................................................................... 51
Homographs vs. Controls ......................................................................................... 58
Spanish Lexical Decision Task .................................................................................. 64
Words vs. Nonwords ................................................................................................. 65
Cognates vs. Noncognates ....................................................................................... 71
Homographs vs. Controls ......................................................................................... 78
Verbal Fluency ........................................................................................................... 85
English Picture Naming ............................................................................................ 86
Individual Difference Measures ................................................................................ 86
Matched Subgroup Comparisons ............................................................................. 86
English Lexical Decision Task .................................................................................. 87
Words vs. Nonwords ................................................................................................. 88
Cognates vs. Noncognates ....................................................................................... 92
Homographs vs. Controls ....................................................................................... 95
LIST OF FIGURES

Figure 1-1: Predicted AoA function according to Birdsong (2005). A similar level of native-like L2 attainment should be found among pre-pubescent learners, to the left of 1, followed by a steep drop in L2 attainment during adolescence, between 2-3 in the figure, and ultimately a homogenously low level of L2 attainment achieved by learners beginning after the offset. .................................................................5

Figure 1-2: Scalp distributions displaying the (L)AN and P600 components in low and high proficiency English speakers from Pakulak and Neville (2010). Within the high proficiency (HP) speakers, the LAN is left-lateralized with little to no negativity spreading to bilateral sites, and the P600 is strong in centro-parietal electrodes. The low proficiency (LP) speakers contrast this pattern by showing a bilateral anterior negativity and a much weaker P600 that is not focalized in centro-parietal sites. ............9

Figure 1-3: Results from Dussias and Sagarra’s (2007) eye-tracking study, showing the opposing patterns of L1 processing between bilinguals with extensive L2 exposure and bilinguals with limited L2 exposure/monolinguals. The Spanish bilinguals with extensive English exposure process low-attachment noun phrases (NP2) faster than high-attachment noun phrases (NP1), as do native English speakers. .........................18

Figure 1-4: Results from Levy et al. (2007). The left graph shows that pictures previously named in Spanish were recalled with less accuracy in English than pictures previously named in English, suggestive of English inhibition while naming the pictures in Spanish. The right graph divides the speakers into low and high proficiency based on L1 and L2 picture naming reaction times, whereby small RT differences are higher proficiency speakers and large RT differences are lower proficiency speakers. The inhibitory result persists for the lower proficiency speakers but not at higher proficiencies. .................................................................20

Figure 2-1: Schematic of the procedure used in the present study. All participants completed session one in English; only the monolinguals performed the English picture naming task in session one. L2 learners returned for a second session to perform the lexical tasks in Spanish and the English picture naming task......................36

Figure 3-1: Accuracy results (d’) for the three groups in the English lexical decision task. Advanced learners were significantly more accurate the beginning learners and monolinguals, with no difference between the latter two groups...............................................42

Figure 3-2: Reaction time results for the three groups in the English lexical decision task. All three groups had similar reaction times, with no significant differences between any groups.................................................................43
Figure 3-3: Reaction time differences for words (red) and nonwords (blue) for the three groups in the English lexical decision task. All groups were faster for words compared to nonwords, but did not differ from each other in reaction times. ..................44

Figure 3-4: Accuracy results for words (red) and nonwords (blue) for the three groups in the English lexical decision task. Despite overall high accuracy rates, participants were marginally more accurate for words than nonwords, but none of the groups differed in accuracy. .................................................................45

Figure 3-5: Mean amplitude values (in mv) from 300-500 ms for English words (blue) and nonwords (red) averaged across the three midline electrodes (Fz, Cz, Pz) for all three groups. Words were significantly less negative than nonwords. ..........................................................46

Figure 3-6: Mean amplitudes (in mv) from 300-500 ms for English words (blue) and nonwords (red) across the four regions of interest for all three groups. Advanced learners had the largest difference between words and nonwords, followed by beginning learners, and then monolinguals. ..............................................................48

Figure 3-7: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -4 to 4 mv) for all three groups showing the difference between English words and nonwords. ..................................................................................50

Figure 3-8: Accuracy rates for cognates (red) and noncognates (blue) for all three groups in the English lexical decision task. Participants were more accurate for cognates than noncognates, but none of the groups differed in overall accuracy rates.............51

Figure 3-9: Mean amplitude values (in mv) from 300-500 ms for English cognates (blue) and noncognates (red) averaged across the three midline electrodes (Fz, Cz, Pz) for all three groups. Cognates were less negative than noncognates prior to the Greenhouse-Geisser correction..................................................................................52

Figure 3-10: Mean amplitudes (in mv) from 300-500 ms for English cognates (blue) and noncognates (red) across the four regions of interest for all three groups. No group interactions were statistically significant, but appear to be present in these data..........54

Figure 3-11: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -1 to 1 mv) for all three groups showing differences between English cognates and noncognates. ..........................................................................................56

Figure 3-12: Accuracy rates for homographs (red) and controls (blue) for all three groups in the English lexical decision task. Participants were more accurate for controls than homographs, including monolinguals.................................................................58

Figure 3-13: Mean amplitude values (in mv) from 300-500 ms for English homographs (blue) and controls (red) averaged across the three midline electrodes (Fz, Cz, Pz) for all three groups. Homographs show different effects in different groups, though not statistically significant........................................................................59
Figure 3-14: Mean amplitudes (in mv) from 300-500 ms for English homographs (blue) and controls (red) across the four regions of interest for all three groups. No group interactions were statistically significant, but beginning learners appear to have reduced negativity for homographs across ROIs whereas monolinguals appear to have increased negativity for homographs across ROIs. Advanced learners show no effect of homograph status. .................................................................61

Figure 3-15: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -1 to 1 mv for beginning, advanced learners, -2 to 2 for monolinguals) for all three groups showing differences between English homographs and controls. ........................................................................63

Figure 3-16: Reaction times to words (red) and nonwords (blue) for beginning and advanced learners in the Spanish lexical decision task. Larger variance for beginning learners in response to words suggests that they may be guessing (quickly) for some correct responses and fully processing the words (more slowly) other times. ...............64

Figure 3-17: Accuracy rates for Spanish words (red) and nonwords (blue) for beginning and advanced learners. Again, the large variance in beginning learners suggests a high rate of guessing. .................................................................65

Figure 3-18: Mean amplitudes (in mv) from 300-500 ms for Spanish words (blue) and nonwords (red) for beginning and advanced learners across midline electrodes (Fz, Cz, Pz). Both groups showed a main effect of word status but did not differ from each other.................................................................66

Figure 3-19: Mean amplitudes (in mv) from 300-500 ms for Spanish words (blue) and nonwords (red) across the four regions of interest for advanced and beginning learners. The main effect of word status is apparent, while there was no effect of group or any group interactions.................................................................68

Figure 3-20: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -2 to 2 mv for beginning leaners, -3 to 3 for advanced learners) for both groups showing differences between Spanish words and nonwords. ..................70

Figure 3-21: Reaction time results for Spanish cognates (red) and noncognates (blue) for advanced and beginning learners. Cognates elicited faster reaction times than noncognates, and advanced learners responded marginally faster than beginning learners. No cognate facilitation is present in the beginning learners, likely due to the large variance and the low accuracy rate, which suggests they were guessing, thereby washing out any effects. .................................................................71

Figure 3-22: Accuracy rates for Spanish cognates (red) and noncognates (blue) for advanced and beginning learners. Participants were more accurate for cognates than noncognates. Advanced learners were more accurate than beginning learners. Though the magnitude of accuracy differences across groups looks comparable, the variance in the beginning group rendered the difference between cognates and noncognates insignificant for that group. .................................................................72
Figure 3-23: Mean amplitudes (in mv) from 300-500 ms for Spanish cognates (blue) and noncognates (red) for beginning and advanced learners across midline electrodes (Fz, Cz, Pz). The main effect of cognate status is driven by the cognate effect present in the advanced learners.

Figure 3-24: Mean amplitudes (in mv) from 300-500 ms for Spanish cognates (blue) and noncognates (red) across the four regions of interest for advanced and beginning learners. Cognates are consistently less negative within the advanced learners but has little to no effect within beginning learners.

Figure 3-25: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -1 to 1 mv for beginning learners, -2 to 2 for advanced learners) for both groups showing differences between Spanish cognates and noncognates.

Figure 3-26: Accuracy rates for Spanish homographs (red) and controls (blue) for advanced and beginning learners. Both groups have higher accuracy rates for homographs compared to controls, but once again, the large variance within the beginning learners’ responses reduced the statistical difference between the two word types.

Figure 3-27: Mean amplitudes (in mv) from 300-500 ms for Spanish homographs (blue) and controls (red) for beginning and advanced learners across midline electrodes (Fz, Cz, Pz). The interaction revealed that beginning learners have a graded effect of homograph status from anterior to posterior sites.

Figure 3-28: Mean amplitudes (in mv) from 300-500 ms for Spanish homographs (blue) and controls (red) across the four regions of interest for advanced and beginning learners. As found in the midline analysis, beginning learners show reduced negativity for homographs in posterior regions and increased negativity in anterior regions. Advanced learners show homograph interference (in the form of increased negativity) across the scalp.

Figure 3-29: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -1 to 1) for both groups showing differences between Spanish homographs and controls.

Figure 3-30: Verbal fluency performance in English (left) and Spanish (right). No groups differed in performance in the English verbal fluency, whereas advanced learners performed better than beginning learners in Spanish.

Figure 3-31: Accuracy (d’) in the English lexical decision for advanced and beginning learners and the matched monolingual subgroups. Advanced learners were more accurate than the matched monolingual group, but the beginning learners and monolinguals did not differ significantly.

Figure 3-32: Accuracy rates for English words (red) and nonwords (blue) for advanced learners and monolinguals (left) and beginning learners and monolinguals (right).
Advanced learners were more accurate than the matched monolinguals. There was not an effect of word status.

Figure 3-33: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -4 to 4) for learners and the matched monolingual groups showing differences between English words and nonwords.

Figure 3-34: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -2 to 2) for learners and matched monolinguals showing differences between English cognates and noncognates.

Figure 3-35: Scalp distribution for the beginning learners from 550-650 ms, revealing a small (~1 mv) effect for cognates in centro-parietal regions consistent with an N400.

Figure 3-36: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -2 to 2) for learners and matched monolinguals showing differences between English homographs and controls.

Figure 3-37: Waveforms for advanced learners and matched monolinguals from an anterior electrode (Fz) showing increased negativity for homographs compared to control words in both groups.

Figure 3-38: Waveforms for beginning learners and matched monolinguals from a right-lateralized posterior electrode (P4), revealing a similar magnitude of reduced negativity for homographs and control words.

Figure 3-39: Relationship between L2 proficiency (L2/L1 verbal fluency measure) and English performance (d’) reveals a positive correlation (r = .38, p = .02). Higher L2 proficiency is related to higher L1 proficiency.

Figure 3-40: Relationship between English cognate effect and English homograph effect (r = -.28, p = .06). Behaviorally, these effects are negatively correlated.

Figure 3-41: Relationship between English cognate effect and English homograph effect on the ERP record shows a positive relationship (r = .3, p = .05).

Figure 3-42: Relationship between L2 sensitivity (English cognate effect) and English performance (English d’). Greater cognate interference is related to lower English performance.

Figure 3-43: Relationship between Spanish proficiency, as operationalized by the word vs. nonword distinction on the ERP record, and L2 sensitivity, as operationalized by the English cognate effect. Greater English cognate facilitation is associated with a larger N400 for Spanish words vs. nonwords.

Figure 3-44: Relationship between behavioral Spanish word effect (word RT – nonword RT) and L2 sensitivity (English cognate effect). Greater facilitation for English cognates is related to a larger difference in reaction times between Spanish words and nonwords.
LIST OF TABLES

Table 1-1: Outline of the tasks used in the study. *Only L2 learners performed tasks in the L2 .......................................................... 28

Table 1-2: Descriptive statistics for all participants, learners grouped by proficiency. Means are given for self-rated proficiency measures, age, individual difference measures of working memory (operation span) and executive function (flanker task), and proficiency in an on-line production task (verbal fluency). Advanced learners rated their L1 proficiency significantly higher than beginning learners, but neither group differed from the monolinguals. Advanced learners rated their L2 proficiency higher than beginning learners, who rated their L2 proficiency higher than monolinguals. Advanced learners also produced significantly more exemplars in the Spanish verbal fluency task than the beginning learners........................................28

Table 1-3: Descriptive statistics on the same measures given in Table 2, but for groups in a matched subset of monolinguals and learners. Both groups of learners rated their L2 proficiency significantly higher than the matched monolinguals. The groups of monolinguals marginally differed in age, but neither differed from the groups of learners. The advanced learners produced more Spanish verbal fluency exemplars than the beginning learners.................................................29

Table 3-1: Results of the 3 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for words and nonwords in the English lexical decision task...............................................................45

Table 3-2: Results of the 2 x 2 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for words and nonwords in the English lexical decision task...............................................................47

Table 3-3: Results of the 3 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for cognates and noncognates in the English lexical decision task...............................................................52

Table 3-4: Results of the 2 x 2 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for cognates and noncognates in the English lexical decision task...............................................................53

Table 3-5: Results of the 3 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for homographs and controls in the English lexical decision task...............................................................58

Table 3-6: Results of the 2 x 2 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for homographs and controls in the English lexical decision task...............................................................60
Table 3-7: Results of the 3 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for words and nonwords in the Spanish lexical decision task. .................................................................66

Table 3-8: Results of the 2 x 2 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for words and nonwords in the Spanish lexical decision task. .................................................................67

Table 3-9: Results of the 3 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for cognates and noncognates in the Spanish lexical decision task. .................................................................73

Table 3-10: Results of the 2 x 2 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for cognates and noncognates in the Spanish lexical decision task. .................................................................74

Table 3-11: Results of the 3 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for homographs and controls in the Spanish lexical decision task. .................................................................79

Table 3-12: Results of the 2 x 2 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for homographs and controls in the Spanish lexical decision task. .................................................................81
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the committee members who helped guide this project and provided valuable feedback in the stages of development, execution, and finalization. I would especially like to thank my advisor, Dr. Judith Kroll, for her wonderful support through the many phases of the research process. Her feedback was critical to the outcome of this project, and the time she dedicated to assisting me was very generous. My lab members, Rhonda McClain and Jason Gullifer, were incredibly helpful in all matters, particularly in analyzing the ERP and behavioral data, and their experience in the lab made them wonderful mentors. Finally, I would like to thank my family, who always encouraged me to set ambitious goals and pursue them with all my heart.
A common observation in research on second language learning is that children are better language learners than adults. For children, language learning appears to come naturally, but this ability declines with age. Adults often struggle for years to learn any second language (L2) to high proficiency. Many past studies have focused on identifying the factors that constrain or facilitate adult L2 learning, leading to debate as to whether native-like proficiency can indeed be achieved past early childhood. While this debate has led to an upsurge in research comparing L2 processing in highly proficient adult bilinguals to monolingual native language (L1) speakers, only recently has research begun to investigate how high proficiency in an L2 produces changes to a bilingual’s native language. Growing evidence indicates that high L2 proficiency may alter L1 usage by slowing lexical access and producing convergence between the L1 and L2, with effects of co-activation evident even when bilinguals use the L1 in an exclusively native language context. Based on this evidence, the present study sought to test a novel contributing factor that influences adult L2 learning, specifically the flexibility of the L1 to accommodate the changes described above. The hypothesis is that L1 changes may be a necessary step during L2 learning, and that adult L2 learners who are better able to tolerate changes to the L1 may be more successful in achieving high proficiency in the L2.

First, the traditional debate on age of acquisition and L2 vs. native-like L1 processing will be reviewed. Following that, I discuss the issue of whether native-like processing can truly serve as a benchmark against which L2 learners can be compared. Given the variability that encompasses L1 processing, the evidence demonstrating that the L1 changes in highly proficient bilinguals is less surprising. I review evidence showing both L1 costs and the sensitivity of the L1
to the L2 in highly proficient bilinguals will be discussed, and discuss the proposed mechanism behind such changes. Finally, the present study will be described in further detail, along with the tasks, predictions, and analyses that were planned a priori.

Age of Acquisition and the Critical Period Hypothesis

Past research investigating the age of acquisition (AoA) has shown that L2 phonological and grammatical access declines with increasing AoA. In the classic studies on this issue, Johnson and Newport (1989) examined the effects of AoA on the comprehension of L2 grammar and Flege, Munro and MacKay (1995) studied how AoA affects the degree of perceived foreign accent in the L2. The results in both studies showed fundamentally the same pattern: a linear decline in L2 acquisition as AoA increased. Someone who begins to learn an L2 later in life will, on average, have a stronger foreign accent and reduced accuracy in discriminating grammatical and ungrammatical sentences. This general pattern of results has been replicated in many studies (see Dekeyser, 2005, for review).

The Critical Period Hypothesis and Shallow Structure Hypothesis

The question about AoA is not whether the data are reliable, but how they should be interpreted. One interpretation of the steep decline with increasing AoA is the maturational account or critical period hypothesis. Supporters of this account (Clahsen & Felser, 2006; Lenneberg, 1967) argue that L2 learning, like many other cognitive functions, has a critical or sensitive period during which the L2 can be successfully learned and after which it is sparsely or minimally learned. AoA has been hypothesized to affect the L2 phonology and grammar more than the lexicon. Many past studies have tested the hypothesis that late L2 learners can process
the lexico-semantics of an L2 but lack the plasticity to detect subtle violations at the level of the grammar.

Indeed, these predictions garnered support from many experiments, including some using event-related potentials (ERPs) to provide a more sensitive test of the hypothesized constraints on L2 processing (Hahne & Friederici, 2001; Weber-Fox & Neville, 1996). In an experiment by Hahne and Friederici (2001), native Japanese speakers who learned German in adulthood and were currently living in Germany made grammaticality judgments on German sentences. Some of the incorrect sentences contained a purely semantic violation, some contained a purely syntactic violation, and others were both semantically and syntactically incorrect. Like native speakers, these participants produced an N400 in response to the semantically incorrect sentences, but differed from native speakers in the responses elicited from the syntactically incorrect sentences. The results further indicated that the responses to syntactic violations in the L2 learners were of a qualitatively rather than quantitatively different nature from native German speakers. The researchers concluded that high proficiency late bilinguals do not rely on the same mechanism or learning processes as native speakers of a language, supporting an interpretation that AoA effects reflect maturational constraints.

What components of language processing might be affected by hard constraints? According to the Shallow Structure Hypothesis (Clahsen & Felser, 2006), adult L2 learners may come to process lexico-semantic aspects of sentences as deeply as native speakers but will always have a shallower understanding of the syntax and thus will never attain full native-like proficiency. Their argument is made through comparisons between child and adult L2 learners, concluding that child learners are quantitatively different than mature speakers with respect to working memory and lexical retrieval mechanisms, but with time will actualize to become native-like speakers. Adult L2 learners, however, are comparable to mature native speakers in terms of working memory and efficiency of retrieval, yet still fail to appreciate the L2 syntax, particularly
in the inability to exploit long distance dependencies. They argue that adult learners rely instead on semantics and pragmatics to compensate for their inability to process the syntax skillfully. This hypothesis has support from experiments like that of Hahne and Frederici (2001) which demonstrate that late L2 learners do not process the L2 similarly to native speakers, though other recent ERP studies have actually found groups of extremely proficient late L2 learners who do appear to be native-like in their processing of L2 sentences (Morgan-Short, Steinhauer, Sanz, & Ullman, 2012).

**Alternative Accounts to the Critical Period Hypothesis**

Not everyone agrees that lack of plasticity is the culprit behind the observed effects of AoA. In a review of the literature, Birdsong (2005) evaluated the specific predictions that the critical period hypothesis would make regarding age effects in L2 acquisition. The critical period account would predict a function that resembles a stretched Z, with a high but level attainment of the L2 in the beginning of life until ~12-13 years (puberty), followed by a steep drop off and a plateau in performance throughout the rest of life. The AoA data from the previous studies do not appear to fit this function, even when the samples are separated into pre- and post-pubescent age groups. Instead, the AoA function is linear and L2 performance never seems to reach plateau or floor performance. In fact, there does not seem to be consistently high L2 performance among young children either; a linear decline is still observed within the pre-pubescent children. Furthermore, the post-pubescent lack of plasticity should apply universally, and thus this cannot account for the individuals who do seem to achieve what some consider native-like performance. Indeed, more recent ERP evidence does indicate that native-like processing can be achieved in late L2 learners (Morgan-Short et al., 2012).
Morgan-Short and colleagues (2012) taught participants an artificial language through either implicit or explicit training methods and tested them at different points throughout the training both behaviorally and with ERPs. Explicit training involved providing the participants with metalinguistic information about the structure and rules of the language along with examples in which the rules were applied, which mimics a classroom learning setting. Implicit training involved the same amount of training time, but let the learners deduce the rules through practice with feedback, which mimics an immersion learning setting. The group that was trained explicitly showed non-native like processing for syntactic violations, similar to many other studies concluding that late L2 learners cannot learn the L2 to a native-like proficiency. However, the implicitly trained group did demonstrate native-like responses to syntactic violations by the second testing session, suggesting that late L2 learners can come to process in a native-like manner and also that the language learning context may be crucial for achieving this.

Figure 1-1: Predicted AoA function according to Birdsong (2005). A similar level of native-like L2 attainment should be found among pre-pubescent learners, to the left of 1, followed by a steep drop in L2 attainment during adolescence, between 2-3 in the figure, and ultimately a homogenously low level of L2 attainment achieved by learners beginning after the offset.
An alternative account to the critical period hypothesis focuses instead on the interactions between L1 and L2 proficiency (e.g., Jia, Aaronson, & Wu, 2002; Jia & Aaronson, 2003). Lower L2 attainment may be due to higher L1 proficiency at the time of L2 learning. Since L1 proficiency increases throughout the lifespan, they hypothesize that achieved L2 proficiency is inversely related to the level of L1 proficiency at the time of learning. This relates directly to L1 entrenchment and stability as age increases (Elman, 1993). Instead of viewing AoA effects as a maturational constraint, this proposal views these effects as a learning phenomenon in which the entrenchment of the L1 constrains L2 learning.

Jia and Aaronson (2003) conducted a longitudinal study to investigate the interaction between the L1 and L2 during a three-year period of L2 immersion. They tested Chinese children who recently immigrated to the United States and ranged from ages 5-16 at the time of the move. The results showed that the children with higher L1 (Chinese) proficiency at the beginning of the study had lower L2 (English) attainment and were less likely to shift language dominance. L1 proficiency was necessarily related to age, as older children have more developed cognitive and language systems, but above and beyond age, the younger children were more likely to attrite in their L1 whereas the older children mostly maintained their L1 performance. Younger children also achieved higher proficiency overall and voluntarily used the L2 in more everyday activities. While the general pattern of AoA remains the same in this experiment, the researchers measured additional variables that allow a more subtle interpretation of the findings. This account bears a resemblance to the hypothesis tested here with regard to the relationship between the L1 and L2, but differs from it in the same way that many other current accounts do: it maintains that the L1 is stable.

The stability of the L1 is an assumption of both AoA accounts described above. Despite the prevailing view that the L1 is stable, the more recent literature reports measurable changes to the L1 in adults who have managed to achieve high proficiency in an L2. Traditionally, language
learning research has operated under the assumption of hard constraints, in both the L1 and the
L2, characterized by the critical period accounts. Yet recent research in all domains, including
language, continues to demonstrate greater plasticity in adulthood than previously assumed. The
hypothesis guiding the present study capitalizes on this broadening concept of continued
plasticity into adulthood by arguing that not only that an L2 can be learned to a high level of
fluency regardless of when learning began, but also that the process of learning an L2 produces
changes to the L1, which has been assumed stable and privileged.

Native vs. Non-Native Language Processing

Many of the studies described here have reported evidence in favor or against the claim
that L2 learners can achieve “native-like” processing. Interestingly, recent research has found that
“native-like” processing does not apply uniformly to a population of speakers (Pakulak &
Neville, 2010; Tanner & Van Hell, 2012). Intuitively this would seem obvious; among speakers
of a language, there exist large proficiency differences. However, at the neural level (using
ERPs), the differences between native-like processing among individuals of different proficiency
levels within a language are striking. These discrepancies call into question the practice of
comparing L2 processing to monolingual, “native-like” L1 processing as done in many of these
studies, especially those concluding that adult L2 learners cannot achieve native-like proficiency
in the L2 (Hahne & Friederici, 2001).

In the traditional story of native-like processing used by many, syntactic errors (e.g.,
phrase-structure violation) elicit an early left anterior negativity (ELAN) around 200 ms post-
stimulus, followed by a positive peak around 600 ms (P600) (Hahne & Friederici, 2001). The L2
learners who did not show these same two components were regarded as non-native-like in their
L2 (Hahne & Friederici, 2001). However, a study by Pakulak and Neville (2010) tested native,
monolingual English speakers who differed in their English proficiency and found that native
speakers vary in how they display this pattern. They recruited participants from various
backgrounds who contrasted in many sociological factors, primarily socioeconomic status (SES),
which in turn is related to education, upbringing (e.g., how much they were read to as children),
and proficiency. Their participants listened to simple, single-clause sentences that were
grammatical or ungrammatical (e.g. Timmy can ride the horse at his* farm/ Timmy can ride the
horse at my farm) and indicated if the sentence was grammatically correct. They also took
measures of the participants’ English proficiency using a battery of vocabulary and grammar tests
to group the participants into high and low proficiency.

When controlling for the individual differences (e.g., SES, working memory, etc.),
Pakulak and Neville (2010) found that different levels of proficiency in English were associated
with very different patterns of brain activity in response to syntactic violations. Low proficiency
(L1) processing was characterized by a sustained, bilateral anterior negativity throughout the
epoch and a P600 in posterior regions. High proficiency (L1) participants instead showed a short,
discrete anterior negativity that was left lateralized, followed by a very large P600 in posterior
regions that extended into some anterior medial sites (or, “native-like” processing according to
Hahne & Friederici, 2001). The P600 in high proficiency speakers was significantly larger in
amplitude than in low proficiency speakers and extended to broader regions on the scalp.
Interestingly, two “indicators” of non-native processing for late L2 learners is a smaller P600 than
monolinguals and a more bilateral distribution in the anterior negativity (Hahne & Friederici,
2001). Therefore, it may be that L2 learners are showing patterns similar to native speakers, just
not to native speakers who are highly educated, coming from a high SES and above-average in
their language proficiency. L2 learners who pattern similarly to the low proficiency L1 speakers
in this study may strive to achieve higher proficiency like that of the high proficiency speakers
reported here, but by no means should they be characterized as processing qualitatively differently than native speakers given their similarity to the lower proficiency L1 speakers.

Another study that calls into question the validity of the “native-like” pattern of brain activity was conducted by Tanner and Van Hell (2014). They tested the hypothesis that native, monolingual speakers of English may utilize different neural substrates to process morphosyntax. They argue that the biphasic pattern of activity (i.e. LAN followed by P600) found in the grand average of a group of monolinguals may actually be the combination of an N400 response in some participants and a P600 response in other participants, making the LAN a spurious result of averaging positivity and negativity at different latencies. Research on the LAN has suggested that

Figure 1-2: Scalp distributions displaying the (L)AN and P600 components in low and high proficiency English speakers from Pakulak and Neville (2010). Within the high proficiency (HP) speakers, the LAN is left-lateralized with little to no negativity spreading to bilateral sites, and the P600 is strong in centro-parietal electrodes. The low proficiency (LP) speakers contrast this pattern by showing a bilateral anterior negativity and a much weaker P600 that is not focalized in centro-parietal sites.
the component is an automatic response that detects a syntactic violation whereas the P600 is involved in reanalysis and repair. Therefore, the presence of the P600 should depend upon the presence of the LAN, since you cannot repair what you did not detect. On the other hand, if the LAN is simply an N400 in disguise, then participants who show earlier negativity should not show a P600, and those who show a P600 should not show the earlier negativity.

Tanner and Van Hell (2014) asked participants (native, monolingual English speakers) to judge a visually presented sentence as “good” or “bad” (if it was ungrammatical, nonsensical, or otherwise anomalous). Critical stimuli contained subject-verb disagreements or verb-tense disagreements and were intermixed with grammatical filler sentences. They computed the grand average across all participants and replicated the common finding of an early left negativity followed by the P600. However, they then calculated each individual’s N400 magnitude in the 300-500ms time window and P600 magnitude in the 500-800 ms time window over centro-parietal electrodes. As predicted, the N400 and P600 magnitudes were negatively correlated across participants, meaning that participants who showed an N400 in response to the violations did not show a P600 and vice versa. Moreover, the majority of participants had a P600 response, which could contribute to why the P600 has been found more reliably across studies whereas the LAN may or may not appear. They suggest that the presence or absence of the “LAN” is predicated upon the proportion of participants with a predominant N400 response.

Importantly, this study further broadens what may be characteristic of native language processing, given that another common finding in proficient L2 learners is the presence of an N400 where the P600 should be. Furthermore, these results can be applied to the Pakulak and Neville (2010) study that found a prolonged, bilateral anterior negativity in the lower proficiencies and a very large, broadly distributed P600 in the higher proficiencies that extended to anterior medial sites. Averaging across the high and low proficiency groups in that study would
wash out many of the anterior medial effects where the anterior negativity and P600 intersect in time and space.

These studies broaden the scope of what can be considered native-like processing and give perspective to future directions of research. Understanding and characterizing the variability of L1 processing in cross-sectional design provides a scope of how native-like processing can manifest, but research has yet to fully identify and characterize how L1 processing can vary over time due to age and context. Understanding how the L1 changes with increases in L2 proficiency as well as what individual differences predict this kind of L1 flexibility is one goal of the current study.

**L1 Changes in Highly Proficient Bilinguals**

Some research has begun to identify and characterize how L1 processing can vary over time as a result of L2 learning. Recent evidence suggests that highly proficient bilinguals do not process their L1 similarly to monolinguals matched in age and education (e.g., Ameel, Malt, Storms, & Van Assche, 2009; Chang, 2013; Dussias & Sagarraga, 2007; Kroll, Michael, Tokowicz, & Dufour, 2002; Lev-Ari & Peperkamp, 2013; Linck, Kroll, & Sunderman, 2009). This evidence can be broadly organized into two types of L1 change: L1 costs and L2 sensitivity. Before detailing the L1 changes and the evidence for each one, it is important to understand the probable mechanism that underlies both, parallel activation, and how it produces the types of changes seen in these highly proficient bilinguals.
Parallel Activation

One of the most robust findings in the field of bilingualism research in the last two decades is that both of a bilingual’s languages are simultaneously active at all times (for review, see Kroll, Dussias, Bogulski, & Kroff, 2012). Intuitively, the native L1 remains active when using the less dominant L2. However, the converse is also true, such that bilinguals using their L1 also demonstrate concurrent L2 activation. Furthermore, parallel activation or cross-language activation, appears to be present in all bilinguals regardless of the languages they speak, including in bilinguals whose languages use different written scripts, as with Chinese and English (Thierry & Wu, 2007), or different modalities, as with American Sign Language (ASL) and English (Morford, Wilkinson, Villwock, Piñar, & Kroll, 2011).

The evidence revealing parallel activation across the bilingual’s two languages comes from a diverse set of tasks and methods. A common technique for measuring the activity of the nontarget language is to compare the processing of language ambiguous cognates and homographs with words that unambiguously belong to one of the bilingual’s two languages alone. Cognates are words that have similar form and meaning across a bilingual’s two languages (e.g., piano in English and Spanish). The convergence in lexical form and meaning for cognates tends to facilitate processing on word recognition and word production tasks (e.g., Costa et al., 2000; Dijkstra, 2005). In contrast, homographs, which have similar lexical form but conflicting meaning across the bilingual’s two languages (e.g., Spanish carpeta means ‘folder,’ not ‘carpet’), typically produce interference in processing. In theory, monolingual speakers should process cognates and homographs no differently than other words that are matched on lexical properties such as word length and frequency. Therefore, when bilinguals demonstrate sensitivity (i.e., facilitation or interference) to language-ambiguous cognates and homographs compared with matched unambiguous control words, we can infer that the heightened activity of their nontarget language
facilitates processing when form and meaning agree (cognates) or that the heightened activity competes with the target language when there is conflict between form and meaning (homographs).

Studies have reported that a bilingual’s nontarget language is active in a range of tasks that require only one of the two languages to be used. Cognate facilitation has been found for isolated word reading in both the L1 (Van Hell & Dijkstra, 2002) and the L2 (e.g., Dijkstra, 2005) as well as for sentence reading in the L1 and L2 (e.g., Schwartz & Kroll, 2006; Titone, Libben, Mercier, Whitford, & Pivneva, 2011). Bilinguals also demonstrate parallel activation of the two languages while listening to speech in either language (e.g., Blumenfeld & Marian, 2011, 2013). Most surprisingly, both languages appear to be active when bilinguals plan to speak in one language alone, even though initiation of speech planning lies within the control of the speaker (for reviews on bilingual language production, see Kroll, Bobb, & Wodniecka, 2006; Kroll & Gollan, 2014).

Evidence for parallel activation comes from converging methodologies, including behavioral and ERP studies, and is confirmed by many replications. As a relatively accepted finding, research has since moved toward understanding how a bilingual refrains from constantly producing words in the nontarget language, especially the L1, given its dominant and automatic activity. Particularly for beginning L2 learners, for whom the intentional use of the L2 should spuriously produce strong L1 activation, it is important to understand the mechanism of language selection.

One of the prominent hypotheses for how language selection occurs is the Inhibitory Control Model (Green, 1998). Green proposes that an inhibitory mechanism suppresses the activation of the nontarget language to the necessary degree in order for the target language to win the selection process. For a beginning L2 learner to speak the L2, inhibition is required to successfully suppress the L1 competitors. For a more balanced bilingual to speak the L2, less
inhibition may be required to successfully suppress the L1 competition, but in order to achieve high proficiency, bilinguals must engage inhibitory control often. In theory, inhibitory control is honed over time with the excessive practice of engaging and disengaging this mechanism through daily language use, producing many of the bilingual advantages in inhibitory control reported in recent findings (for review, see Bialystok, Craik, Green, & Gollan, 2009).

What remains to be investigated is whether a high level of inhibitory control is necessary in order to learn an L2, if L2 learning improves/influences inhibitory control, or if some combination of these factors produces the bilingual advantage in inhibitory control. The present study attempts to test this by administering tasks that measure cognitive resources (both working memory and inhibitory control) to evaluate the relationship between inhibitory control and L1/L2 proficiency. The prediction is that individual differences in inhibitory control should modulate the changes seen in the L1 and the rate of L2 acquisition, such that adults who begin L2 learning with a higher level of inhibitory control should show greater L1 change and higher L2 proficiency compared to peers with a similar level of L2 exposure. Past research also suggests that higher working memory is related to higher attainment of L2 proficiency (Miyake & Friedman, 1998).

**Influence of the L2 on the L1**

As stated, recent evidence suggests that highly proficient bilinguals differ from monolinguals in L1 processing. One way in which they differ is what the present study refers to as L2 sensitivity. Unlike monolinguals, proficient bilinguals demonstrate sensitivity to the L2 during tasks that only require the L1 to be used. Parallel activation tells us that when using the L1, the L2 representation is simultaneously active. This type of L1 change refers to varying the usage of the L1 due to conscious or unconscious awareness and influence of the L2.
One manifestation of this sensitivity was already mentioned in the section on parallel activation: cognate and homograph effects. For example, Van Hell & Dijkstra (2002) recruited Dutch-English-French trilinguals, concealing the fact that they were recruited for their trilingual status, to perform a lexical decision task in Dutch only. Dutch words and nonwords (or pseudowords) were presented and participants judged whether each item was a real Dutch word or not. Included among the words were English or French cognates. They found that participants who were highly proficient in English but not French were faster to respond to Dutch-English cognates compared to other Dutch words matched on length and frequency, but were no faster to respond to Dutch-French cognates. Participants who were highly proficient in all three languages, on the other hand, were faster to respond to both Dutch-English and Dutch-French cognates. These results suggest that the second (and third) language come to influence how the L1 is processed (in this case, producing facilitation), and furthermore that the degree of influence is related to proficiency in the nontarget language. Cognate effects in the L1 have been replicated and additionally found using ERPs. Midgley, Holcomb & Grainger (2011) found that L1-L2 cognate words elicit a smaller N400 compared to L1 words matched on lexical properties. It therefore appears that highly proficient bilinguals demonstrate sensitivity to words that overlap in the L1 and L2 both behaviorally and neurophysiologically.

Another example of the mutual influence of two languages on each other is called lexical or semantic convergence, described by Ameel et al. (2005; 2009). They studied early bilinguals, or people who were raised from birth knowing two languages. They asked participants to categorize everyday items that have different translations across languages, such as bowls, cups, glasses, and jars. Rather than categorize items as monolinguals of whichever language the bilinguals were using, they found that the bilinguals had merged their lexico-semantic representations to produce naming categories that approximated monolinguals of both languages but matched neither exactly. In other words, simply knowing that the nontarget language would
apply a different term to an object changed the way that the bilinguals named categories in the target language.

The studies reviewed refer to lexical-level changes to the L1. However, L2 sensitivity has been found at all levels of processing, including phonology and syntax. At the phonological level, Chang (2013) compared experienced (proficient) and inexperienced (novice) learners of Korean who were part of an intensive six-week Korean immersion program with monolingual speakers of English. He measured the voice-onset time (VOT) for voiceless stops (/p/, /t/, /k/) in English (L1) at five time points throughout the immersion program (once at the end of the first five weeks). While voiced stops (/b/, /d/, /g/) are similar in VOT across English and Korean, the voiceless stops have a longer VOT in Korean than English. As expected, the English VOT for voiced stops did not differ between groups or across the time points. Interestingly, he found that both immersed groups had longer English VOTs for voiceless stops compared to the English monolinguals, and furthermore that the inexperienced/novice learners showed the greatest degree of shift across time, more so than the experienced learners. This indicates that, at least in an immersion context, knowledge of an L2 may come to affect how a bilingual pronounces the L1, and that the beginning stages of learning may exert a larger influence on changes in the L1 in some cases.

Bilinguals also differ in L1 processing at the level of the syntax. Dussias & Sagarra (2007) investigated how highly proficient bilinguals interpret ambiguous relative clauses (e.g., “the sister [NP1] of the man [NP2] who was ill”- Who was ill?). They created sentences in which the final modifying adjective referred either to the first noun in the relative clause (NP1) or the second noun in the relative clause (NP2) by manipulating the grammatical gender of the modifying adjective (e.g., “El policía arrestó a la hermana del criado que estaba enferma desde hacía tiempo,” meaning The police arrested the sister (fem) of the servant (masc) who had been ill (fem) for a while). Speakers of different languages tend to prefer one interpretation to the other.
when the interpretation is underspecified. For example, a Spanish speaker will typically assume that the sister was ill if grammatical gender cannot disambiguate the context, whereas an English speaker will prefer to interpret that the servant was ill.

They tested native Spanish speakers who had been living in the USA for an extended period of time or a limited amount of time, as well as native Spanish speakers living in Spain. Participants read the Spanish (L1) sentences in which grammatical gender forced one interpretation over the other while their eye movements were recorded. They found that the monolingual speakers in Spain and the bilingual speakers who had been in the USA for a limited amount of time processed the sentences as expected; when grammatical gender forced participants to attribute the modifying adjective to the NP2, these participants showed longer total reading times. However, the Spanish speakers living in the USA for an extended period of time showed the opposite pattern; they required longer total reading times for sentences in which the modifying adjective referred to the NP1, the preferred interpretation for typical Spanish speakers. Their results suggest that sufficient L2 exposure, particularly immersion experience, may shift syntactic processing in bilingual speakers such that they come to process their L1 through the lens of the L2 in certain cases.
As described here, growing evidence suggests that highly proficient bilinguals may no longer be exactly native-like in their own native language. Knowledge of an L2 remains active in the minds of highly proficient bilinguals, influencing how they process the L1. In many cases, such as cognate effects, this L2 sensitivity produces facilitation, and in others, such as lexico-semantic convergence, it simply changes the way in which the L1 is used. There are, however, other instances in which highly proficient bilinguals demonstrate an overall slowing in the L1 compared to monolinguals.

**L1 Costs**

Bilinguals often demonstrate an advantage on tasks measuring inhibitory control (e.g., Abutalebi et al., 2012; Engel de Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok, 2012) and novel language learning (e.g., Kaushanskaya & Marian, 2009), and show facilitation for certain words that have L1-L2 overlap, but not all changes that occur in a bilingual are necessarily
advantageous. In many tasks measuring lexical access, bilinguals appear to be overall slower and have generally less access to the L1 lexicon. This evidence comes primarily from lexical production studies (e.g., Baus, Costa, & Carreiras, 2013; Sandoval, Gollan, Ferreira, & Salmon, 2010). Whereas the evidence for L2 sensitivity was found almost exclusively in highly proficient bilinguals (but see Chang, 2013), evidence for L1 costs can be observed in L2 learners at various proficiency levels.

At low proficiency levels, these L1 costs are perhaps more widespread. Kroll, Michael, Tokowicz and Dufour (2002) asked high and low proficiency L2 learners to perform a variety of tasks, including a measure of working memory and an English (L1) word naming task. Though not the purpose of the study, they unexpectedly found that the low proficiency learners were overall slower than the high proficiency learners to name the L1 words out loud. Considering that the high proficiency learners may have undergone self-selection pressures to only include learners with high levels of cognitive resources to devote to L2 learning, they divided the groups into high and low working memory scores. Even then, the low proficiency learners who did not differ from the high proficiency learners in working memory demonstrated significantly slower L1 word naming latencies. The results suggest that the L2 learning process may incur certain costs to L1 lexical access, producing an overall slowing.

Another study using a lexical production task by Levy et al. (2007) found similar results. They asked English learners of Spanish to perform an interleaved picture-naming task, in which some pictures were named in English and some in Spanish. Each picture could be named up to ten times, but every time in the same language as it was previously named. Following that, participants performed a final test in English (L1) in which they were provided a rhyming cue word paired with the first letter of the target picture (e.g., break-s____) and asked to produce the name of the picture from the previous task (e.g., snake). As expected, pictures that were named in English more often in the picture-naming task were also correctly recalled more often in the final
test. However, pictures that were named in Spanish more often in the picture-naming task showed the opposite pattern: a picture that was previously named once in Spanish had a higher rate of recall in English than a picture that was previously named ten times in Spanish. They then used the picture-naming reaction times in each language to divide the groups into high and low proficiency, such that a smaller difference in English and Spanish RTs in the picture-naming task corresponded to higher proficiency. This pattern persisted for the lower proficiency learners but not for the higher proficiency learners, again confirming the results of the Kroll et al. (2002) study showing these costs primarily during the learning process.

Both studies described thus far compared low and high proficiency L2 learners, but how do bilinguals compare to monolinguals in L1 production tasks? Sandoval, Gollan, Ferreira and Salmon (2010) asked English dominant bilinguals and English monolinguals to perform a verbal fluency task for both semantic (e.g., animals) and letter (e.g., words that start with the letter s)
categories. They measured the total number of correct responses, first response latency, and the mean subsequent response latency (i.e. *fulcrum point*, thought to measure lexical retrieval).

Compared to monolinguals, they found that bilinguals produced fewer correct responses, longer first response latency, and delayed retrieval as indicated by a longer fulcrum point. Bilinguals also produced lower frequency words on average and a higher proportion of cognate words compared to monolinguals. Their results support with the interference account that proposes that the delayed and reduced lexical access in bilinguals is due to nontarget language interference, since bilinguals must manage the interference from the nontarget language in order to retrieve and produce the target language. Their results are also confirmed by many other studies using verbal fluency (Baus et al., 2013; Gollan, Montoya, & Werner, 2002; Linck et al., 2009).

At face value, these L1 costs would appear to be disadvantageous in every sense. However, parallel activation and the Inhibitory control Model (Green, 1998) may help to understand these L1 costs in a different light. Recent research has found that inhibiting the L1 while speaking the L2 has persistent and long-lasting effects (Linck et al., 2009; Misra, Guo, Bobb, & Kroll, 2012). It may be that these L1 costs are the residual results of managing the spurious L1 activation present during L2 use. Modulating access to the L1 in exclusively L1 contexts may ultimately promote inhibitory control and L2 learning by increasing access to the L2, despite the small tradeoff in L1 lexical access. Furthermore, the scope of the L1 costs appears to be limited to the language domain, rather than overall cognition, as evidenced by the fact that bilinguals are overall faster on inhibitory control tasks, show smaller inhibitory effects (Costa, Hernández, & Sebastián-Gallés, 2008), and working memory size does not modulate the costs found at lower proficiencies (Kroll et al., 2002).

While the concept of incurring costs to the native language during L2 learning may seem strange, it would not be the only instance in which cognitive costs during training/learning lead to better retention. In fact, research in the domain of memory and learning developed the concept of
desirable difficulty (Bjork, Bjork, & McDaniel, 2011). Desirable difficulty refers to difficulty induced during training or learning that promotes the encoding, storage, or retrieval of a to-be-learned item.

A recent study testing conditions of desirable difficulty included vocabulary learning as the outcome to be measured. Healy and Bourne (2013) trained English speakers on French vocabulary, interleaving training and testing three times during the first session. In the testing portion, half the participants were provided the French word and asked to produce the English translation (easy condition), whereas the other half were provided the English word and produced the French translation (hard condition). Language research on L2 vocabulary acquisition provided the basis for the easy and hard conditions, finding that L1 to L2 translations are harder than L2 to L1 translations (Kroll & Stewart, 1994). Participants then completed an immediate retention test, and returned one week later for a delayed retention test. They found that the participants who were provided the English cue and asked to produce French (hard condition) did not perform as well as the other (easy) group during the immediate retention test, but maintained their performance after a week (therefore exceeding the performance of the other group after one week), whereas the group with easier conditions had a significant decline in accuracy between the first and second session.

Taken together, evidence suggests that L2 learning may incur a cost to the L1, but the costs may ultimately enhance L2 learning. Given how automatic and reliable the native language is, and has always been, for late L2 learners, tolerating changes and costs to the L1 may be easier for some L2 learners than others. Although not necessarily a conscious or even noticeable decision, it may be that individual differences in cognitive resources or inhibitory control modulate the severity of the costs or the tolerance to these changes.
The Present Study

To review, many have argued that L2 learning has a critical or sensitive period, after which L2 learning will never actualize completely. The evidence for this comes from studies comparing proficient bilinguals to monolinguals showing different patterns of ERP components, concluding that bilinguals are qualitatively different and will never have full access to the L2 grammar. However, other research has characterized the variability within monolingual language processing, inadvertently demonstrating how the patterns produced by proficient bilinguals in the L2 may be “native-like” in many ways. Rather than focusing on native vs. non-native processing, the present study sought to characterize the scope of differences within L1 processing between L2 learners and monolinguals, and within L2 learners at different levels of skill, to understand the trajectory of how changes in L1 processing are related to L2 proficiency.

Despite differences in motivation, presumably the goal of L2 learning is to acquire L2 proficiency. In a classroom learning setting, an unconscious assumption is that students begin the L2 learning process with comparable L1 proficiency, and that the L1 remains stable and relatively unchanged throughout the acquisition process. However, as any teacher can attest, there are clear L1 proficiency differences even among a fairly homogenous group of adult university students, and as recent research has shown, the L1 of a proficient bilingual does appear to be different than the L1 of a monolingual in certain ways. Rather than focusing on the L2 of adult learners, the present study shifted the focus to the L1, seeking to understand how a priori individual L1 differences are related to the differences other researchers have found in the proficient bilinguals’ L1.

The L1 differences found in proficient bilinguals manifest in primarily two ways. Successful, highly proficient L2 learners differ in their L1 by revealing the influence of the L2 during unilingual L1 processing. Evidence for this comes in the form of cognate or homograph
effects in the L1, revealing the strength of the concurrent L2 activation. Other studies have demonstrated similar changes to the L1 syntax (Dussias & Sagarra, 2007) and phonology (Chang, 2013). This L1 change is almost exclusively found in highly proficient bilinguals, with little evidence of such changes in the beginning stages of learning (with the exception of Chang, 2013). It may be that the parallel L2 activation must reach a threshold before exerting influence on the L1, and/or that individuals differ in the threshold required for the influence of the L2 to manifest in L1 processing. Therefore, the present study also utilizes ERPs to provide a more sensitive measure of the neural processes underlying L1 use, given that past research has found neural sensitivity to learning before evidence of that learning has appeared in behavioral measures (McLaughlin, Osterhout, & Kim, 2004).

The other reported L1 difference in proficient bilinguals comes in the form of L1 costs, resulting in slower reaction times and reduced lexical access. Unlike the influence of the L2 on the L1, this type of L1 change is also found in less proficient L2 speakers, particularly those in the process of L2 learning. This trajectory suggests that the costs may be a learning phenomenon related to the modulation and inhibition of the L1 as a means to increase access to the L2 and promote L2 learning. Or, more provocatively, it may be the result of managing a new type of competition for the L1, which has never previously faced competition from a new language.

With a new focus on the L1 during L2 learning, three things become apparent: learners differ from each other in the L1 prior to L2 learning, learners may endure costs to the L1 in the process of L2 learning, and successful L2 learners reveal an influence of the L2 during unilingual L1 processing. Together, this suggests that individual differences in the L1 may predict L2 outcomes, particularly the ability of the L1 to change in response to different demands placed upon the linguistic system during L2 learning that enable L2 accommodation. The hypothesis tested here is that L2 learners whose native language is more tolerant to change will be more successful in achieving L2 proficiency.
The learner’s already established language processing pathways must be open and able to accommodate a new, interdependent language system. In the initial stages of L2 learning, this accommodation is disruptive to the efficiency of the established L1 system, resulting in costs to the L1. Higher degrees of integration result in more strongly coupled L1-L2 activation, or parallel activation, which in turn allows for the potential influence of the L2 on the L1, or L2 sensitivity. However, individuals may differ in the rate and degree of integration, which may predict or be the result of L2 proficiency. Therefore, tolerance to change was evaluated by seeking and measuring L1 costs and L2 sensitivity in a group of L2 learners varying in L2 proficiency, comparing their L2 performance to the other learners, and comparing L1 performance to a group of monolinguals.

The strongest and ideal test of this hypothesis would utilize a longitudinal design, tracing L1 changes within an individual across time. To justify such an approach, this study was first conducted in a cross-sectional design. To investigate L1 costs, a subset of L2 learners (including beginning and advanced learners) were compared to matched monolingual controls in their L1 comprehension and production abilities. Evidence for the influence of the L2 on the L1 was sought in the form of cognate or homograph effects in the comprehension task, as has been done in past research (Van Hell & Dijkstra, 2002), but included ERP sensitivity to cognates and homographs in addition to behavioral effects (McLaughlin et al., 2004; Midgley et al., 2011).

An additional variable of interest is the role of executive function, and more specifically inhibitory control. Looking once again to the literature on proficient bilinguals, many have reported a bilingual advantage in tasks that require inhibitory control compared to monolinguals (for review, see Bialystok & Craik, 2010). This may be a self-selection effect, or it may be that inhibitory skills are honed over time with the constant practice of managing linguistic competition between languages that are active in parallel, or it may be an interaction between the two. However, this finding is pertinent for the present study, given the focus on individual differences and especially given the potential explanation that L1 costs are due in part to the
requirement to inhibit the L1 during L2 use. Therefore, two tasks measuring inhibitory control were included here, under the premise that inhibitory control may modulate the magnitude or type of L1 change found in L2 learners.

A measure of working memory is also included, based on a plethora of research that finds strong correlations between working memory capacity and L2 proficiency (for a review, see Miyake & Friedman, 1998). A larger working memory capacity allows L2 learners to more easily evaluate and connect new L2 words and structures to concepts and L1 translations, which leads to a stronger long-term representation. Interestingly, the correlation between working memory and L2 proficiency is stronger within already proficient L2 learners than within beginning L2 learners (Hulstijn & Bossers, 1992), suggesting that working memory may play a different role in language acquisition in different phases.

Finally, an interesting and quite counterintuitive prediction that the current hypothesis makes is regarding L1 transfer into the L2. Transfer is the process of using L1 knowledge to aid L2 acquisition. L2 learners often transfer any “pertinent” L1 knowledge to guide L2 learning. Pertinent is a relative term, since L1 transfer can both help and hinder the learner at times, depending on the structures present or absent in one or both languages. Structures that are present and similar in both languages are helpful to transfer; for example, the transfer of subject-verb-object word order from English to Spanish is helpful since both languages implement this word order in a majority of situations (though Spanish offers a wider variety of word order combinations). Structures that are absent in the L1 but present in the L2 tend to be very difficult for adult learners to acquire; for example, an English learner of Spanish often struggles with grammatical gender (e.g., Arnon & Ramscar, 2012; Grüter, Lew-Williams, & Fernald, 2012). There is still debate about how structures that are present in both languages but differ in fundamental ways affects L1 transfer (Tokowicz & MacWhinney, 2005). For example, does having Spanish as an L1 help or hinder learning of German grammatical gender, which employs
three genders (masculine, feminine, neuter) that often conflict with an object’s grammatical gender in Spanish? The L1 knowledge that is helpful or hurtful may ultimately depend on the combination of languages and structures being discussed.

Models of language vocabulary learning, such as the Revised Hierarchical Model (Kroll & Stewart, 1994) illuminate the direct role of the L1 in L2 learning through transfer. In this model, newly acquired L2 words are linked to the L1 translation equivalent via lexical links, and gain direct access to the underlying concept as proficiency increases. The current hypothesis would predict that learners who are able to minimize the role of the L1 and manage L1 interference in this process would gain conceptual access faster, and ultimately attain L2 proficiency more rapidly and easily. This prediction can be evaluated by looking at the L2 data for L1 sensitivity, which is the same as L2 sensitivity but in the opposite direction. Therefore, successful beginning L2 learners should demonstrate negligible or smaller cognate effects in the L2, both behaviorally and on the ERP record, compared to L2 learners who are less tolerant to L1 change. This finding would indicate that the activity of the L1 was reduced during L2 use, a difficult feat for even highly experienced bilinguals.

Overview of Tasks, Predictions, Planned Analyses

The studies reviewed above suggests that there are language changes at all levels of language processing, though the current study has limited the scope to the level of the lexicon with plans in the future to extend the investigation to the levels of phonology and grammar. Therefore, the set of tasks were designed to evaluate both the L1 and the L2 in comprehension and production at the lexical level, as well as individual differences in cognitive measures of working memory and inhibitory control. Before describing each of the tasks and procedure in detail (see methods section below), I describe the general logic of the design.
Table 1-1: Outline of the tasks used in the study. *Only L2 learners performed tasks in L2

<table>
<thead>
<tr>
<th>Lexical Tasks</th>
<th>Comprehension</th>
<th>L1</th>
<th>L2*</th>
<th>Behavioral and ERPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>Verbal Fluency Task</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Decision Task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measures of Cognitive Resources</td>
<td>Working Memory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Operation Span (O-Span) Task</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inhibitory Control</td>
<td>Flanker Task AX-CPT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1-2: Descriptive statistics for all participants, learners grouped by proficiency. Means are given for self-rated proficiency measures, age, individual difference measures of working memory (operation span) and executive function (flanker task), and proficiency in an on-line production task (verbal fluency). Advanced learners rated their L1 proficiency significantly higher than beginning learners, but neither group differed from the monolinguals. Advanced learners rated their L2 proficiency higher than beginning learners, who rated their L2 proficiency higher than monolinguals. Advanced learners also produced significantly more exemplars in the Spanish verbal fluency task than the beginning learners.

<table>
<thead>
<tr>
<th></th>
<th>Advanced Learners</th>
<th>Beginning Learners</th>
<th>Monolinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (# Female)</td>
<td>29 (21)</td>
<td>18 (9)</td>
<td>25 (15)</td>
</tr>
<tr>
<td>L1 Self-rating (sd) Scale: 1-10</td>
<td>9.9 (0.4)</td>
<td>9.1 (0.9)</td>
<td>9.4 (0.6)</td>
</tr>
<tr>
<td>L2 Self-rating (sd) Scale: 1-10</td>
<td>6.9 (1.05)</td>
<td>3.9 (0.9)</td>
<td>2.2 (1.7)</td>
</tr>
<tr>
<td>Age in years (sd)</td>
<td>20.2 (2.3)</td>
<td>20.9 (2.2)</td>
<td>20.4 (2.7)</td>
</tr>
<tr>
<td>O-Span (sd) Number of correct items recalled</td>
<td>45.3 (10.02)</td>
<td>42.4 (14.5)</td>
<td>44.5 (9.3)</td>
</tr>
<tr>
<td>Flanker Effect (ms) Incongruent RT – Congruent RT</td>
<td>49.6</td>
<td>57.5</td>
<td>41.3</td>
</tr>
<tr>
<td>Spanish Verbal Fluency (sd) Number of exemplars</td>
<td>7.9 (3.2)</td>
<td>5.3 (3.9)</td>
<td>NA</td>
</tr>
</tbody>
</table>
The comprehension task was a lexical decision task in which participants judged if an item is a real word in the target language. All L2 learners completed this task in English and in Spanish, whereas monolinguals only completed it in English. In both languages, the set of words contained identical and non-identical cognates as well as homographs between English and Spanish. In this task we also recorded ERPs that are time-locked to the presentation of each item. The prediction is that L2 learners will have overall slower reaction times in English compared to monolinguals (L1 costs) and some L2 learners will demonstrate a reduced N400 in response to English-Spanish cognates compared to matched words (L2 sensitivity) (Midgley et al., 2011).

Table 1-3: Descriptive statistics on the same measures given in Table 2, but for groups in a matched subset of monolinguals and learners. Both groups of learners rated their L2 proficiency significantly higher than the matched monolinguals. The groups of monolinguals marginally differed in age, but neither differed from the groups of learners. The advanced learners produced more Spanish verbal fluency exemplars than the beginning learners.

<table>
<thead>
<tr>
<th></th>
<th>Advanced Learners</th>
<th>Monolinguals matched to Advanced Learners</th>
<th>Beginning Learners</th>
<th>Monolinguals matched to Beginning Learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (# Female)</td>
<td>11 (7)</td>
<td>11 (7)</td>
<td>11 (6)</td>
<td>11 (6)</td>
</tr>
<tr>
<td>L1 Self-rating (sd)</td>
<td>9.7 (0.6)</td>
<td>9.2 (0.7)</td>
<td>9.3 (0.5)</td>
<td>9.7 (0.4)</td>
</tr>
<tr>
<td>Scale: 1-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2 Self-rating (sd)</td>
<td>6.8 (1)</td>
<td>2.5 (1.6)</td>
<td>4.1 (0.8)</td>
<td>1.6 (1.1)</td>
</tr>
<tr>
<td>Scale: 1-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in years (sd)</td>
<td>20.6 (3.1)</td>
<td>19.5 (2.3)</td>
<td>20.4 (1.4)</td>
<td>21 (2.2)</td>
</tr>
<tr>
<td>O-Span (sd)</td>
<td>44.2 (9)</td>
<td>43.3 (8.6)</td>
<td>46.9 (7.8)</td>
<td>48.2 (7.5)</td>
</tr>
<tr>
<td>Number of correct items recalled</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flanker Effect (ms)</td>
<td>44.3</td>
<td>38.8</td>
<td>48.9</td>
<td>47.4</td>
</tr>
<tr>
<td>Incongruent RT – Congruent RT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish Verbal Fluency (sd)</td>
<td>8.3 (1.7)</td>
<td>NA</td>
<td>4.8 (3.2)</td>
<td>NA</td>
</tr>
<tr>
<td>Number of exemplars</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Those learners who demonstrate L1 costs and/or L2 sensitivity should also score higher on L2 proficiency measures.

The lexical production task was a semantic fluency task in which participants must list as many exemplars of a given category that they can think of within one minute, for five different categories. Again, L2 learners performed this task in both English and Spanish. The prediction is that L2 learners who produce fewer English exemplars compared to matched monolinguals will also produce a greater number of Spanish exemplars (or will otherwise demonstrate higher L2 proficiency).

Two tasks were used to measure inhibitory control: Flanker and AX-CPT. The Flanker task has been used by many (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Costa et al., 2008), while the AX-CPT is relatively new but measures two components of inhibition: reactive vs. proactive (Braver, 2012; Morales, Gómez-Ariza, & Bajo, 2013). Both tasks produce dependent variables of reaction time and accuracy, and an inhibitory “score” is typically calculated by subtracting a baseline/control condition from the condition in which inhibitory control is required. Past research has found that bilinguals are overall faster (Costa et al., 2008) in addition to showing smaller (i.e. better) inhibitory scores.

Give the large number of potential comparisons in the dataset, the data analysis strategy for evaluating the predictions outlined above was as follows:

1) Data were compared first at the group level, evaluating any differences among the groups of learners in the L1 and the L2, as well as differences between the groups of learners and monolinguals in the L1. For the latter comparison, a subset of learners were matched to the monolinguals on demographic and cognitive characteristics, including age, gender, working memory, and inhibitory control, to investigate if there are differences between beginning learners and monolinguals in the L1, or between advanced learners and monolinguals in the L1. Group analyses were conducted on behavioral measures of
reaction time and accuracy in the English (L1) lexical decision task as well as the production levels in the English verbal fluency task. ERP data were analyzed for the English and Spanish lexical decision task, comparing neural responses to cognates compared to noncognates, homographs to matched control words, and words compared to nonwords.

2) After group comparisons, the data analysis strategy shifted toward individual differences among L2 learners in L1 and L2 processing and cognitive resources. Continuous measures of L2 proficiency were developed to replace the original L2 groupings and were used to evaluate the relationship between L1 and L2 processing, as well as individual differences in inhibitory control and working memory. Individual differences in ERPs in the L1 and L2 were also used by calculating the mean amplitude of the difference between two conditions (cognate vs. noncognates, homograph vs. matched controls, words vs. nonwords) between 300-500ms post-stimulus over centro-parietal electrode sites (C3, C4, Cz, CP3, CP4, P3, P4, Pz). These values were used in the regressions evaluating individual differences in language processing, proficiency, and cognitive resources.

To summarize, beginning and advanced learners of Spanish will be assessed in the L1 and L2 and compared to each other as well as matched monolinguals in the L1. The advanced learners should be faster and more accurate in the L2 tasks compared to the beginning learners, and the goal is to evaluate if this has any consequences for the L1, and/or if the process of L2 acquisition has consequences for the L1, even at lower proficiencies. The hypothesis is that learners who are more tolerant to changes in their native language will achieve higher L2 proficiency.
Chapter 2

Methods

Participants

In total, 72 participants were tested, including 25 (15 female) functionally monolingual English speakers and 47 (30 female) native English speakers in the process of learning Spanish (hereafter “L2 learners”). Both groups were recruited from classes at The Pennsylvania State University; the L2 learners were from Spanish classes at all levels. The L2 learners were originally recruited to fit into three groups (beginning, intermediate, advanced learners) based on how many semesters of Spanish classes they had taken and whether they had studied abroad. After data collection terminated, the L2 learners were regrouped into beginning and advanced learners based on their self-rated L2 proficiency (Advanced learners’ L2 self-rated proficiency > 5, Beginning learners’ L2 self-rated proficiency < 5). Table 2 summarizes characteristics of each group. Table 3 summarizes the characteristics of a subset of participants matched closely on demographic and cognitive factors for group comparisons between monolinguals and learners. Heritage speakers of Spanish and participants with extensive knowledge of a language other than Spanish were excluded, as assessed by a language history questionnaire.

Materials

The language history questionnaire contained demographic questions in addition to questions about other languages, language use, and self-rated proficiency (on a scale from 1-10,
where 1 indicates no knowledge and 10 indicates native fluency). For the full questionnaire, see Appendix A.

The operation span (O-Span) task (Turner & Engle, 1989) is designed to measure working memory. Participants are presented with a math problem (e.g., $2 \times 5 + 1 = 10$), and they must indicate if it is correct. Upon signaling whether the math problem is correct, they see a single word appear on the screen, which they are instructed to remember and recall at a later point in time. In the beginning of the task, they are prompted to recall the previously presented words after two trials (i.e. two words), but by the end of the task they are asked to maintain and recall up to seven words at a time. The score they receive in this task is the number of items correctly recalled after having correctly solved the preceding math problem.

The AX-CPT task measures reactive and proactive inhibitory control (Braver, 2012; Morales et al., 2013). In this task, participants view a series of five letters, of which the first and last letters are distinguished in red font. They must press the “No” button in response to the first four letters regardless of the letter identity. For the fifth letter, they must press “Yes” if and only if the first letter was an A and the fifth letter is an X. In all other cases, they must press “No” in response to the fifth letter. However, 70% of the trials require a “Yes” response to the fifth letter (AX trials), priming participants to press “Yes”. On 10% of the trials (BX trials), the first letter is not an A, so participants know before reaching the fifth letter that they should press “No”, and therefore can utilize proactive inhibition upon seeing the X in the fifth position. On another 10% of the trials (AY trials), the first letter is an A, priming participants to press “Yes” upon presentation of the fifth letter, but the fifth letter is not an X. These trials impose reactive inhibitory control, since participants must react to the fifth letter quickly and effectively to provide the correct response. Finally, 10% of the trials are BY trials, meaning that they contain neither an A in the first position nor an X in the fifth position, eliminating the need for any type of inhibitory control.
The other task of inhibitory control was a Flanker task. In this task, participants see an arrow in the center of the screen (< or >) and must press the button corresponding to the direction the arrow points. Sometimes, the central arrow is flanked by congruent arrows (e.g., > > > > >), and other times by incongruent arrows (e.g., < < > < <). For a set of trials (go/no-go), the central arrow is flanked by diamonds or Xs, indicating that they must provide a response or withhold from responding. The Flanker effect is calculated by subtracting the mean correct response time for congruent trials from the mean correct response time for incongruent trials.

The production task was a semantic fluency task. Participants are provided with a category and must name as many exemplars of the given category within a minute. The categories used in this study were fruits, vegetables, animals, clothing, and body parts, and colors was the practice category. These categories remained the same across sessions, but sessions were one week apart and L2 learners were unaware that they would be performing the same task in the second session (in Spanish). The categories were chosen due to the vast number of potential exemplars, which would help induce variance within the English version, and because they are quite basic categories that beginning L2 learners begin to acquire in the first or second semester, so even the beginning learners could provide some answers in the Spanish version. Responses were coded as correct if they conceivably fit within the category. In the cases where participants named a superordinate exemplar and subordinate exemplars (e.g., “bird” followed by “dove”, “pigeon”, “hawk”), the superordinate category was not credited but the subordinate items were each credited (Sandoval et al., 2010). For the Spanish version, responses must be within one phoneme of the correct/intended exemplar to be credited. This gave leeway for grammatical gender mistakes if the rest of the word was pronounced correctly (e.g., “zapatas” instead of “zapatos”), but also applied more broadly (e.g., “ceballas” instead of “cebollas”).

The comprehension task was a lexical decision task in which participants saw an item on the screen and indicated if it was a real word in the target language or not. Half the items were
nonwords and half were real words. In each language, the real words included 18 nonidentical
cognates (e.g., English pure, Spanish puro), 18 identical cognates (e.g., superior), and 38
homographs (e.g., English exit, Spanish éxito, meaning success). For each group of critical words,
sets of language unambiguous control words were matched on length and frequency. In English,
frequency measures were taken from Francis & Kucera’s (1982) database using the English
Lexicon Project (Balota et al., 2007), which also provided mean lexical decision times for each
word, another parameter used in creating matched lists. In Spanish, frequency measures were
taken from the Alameda & Cuetos (1995) database. The nonwords were matched on length to the
real words. English nonwords were generated using the English Lexicon Project, and Spanish
nonwords were created by swapping a couple letters from a real Spanish word (while still
satisfying phonotactic constraints in Spanish). The stimuli were counterbalanced across sessions
so that each participant only saw a word once and in one language. See Appendix B for the lists
of stimuli and lexical properties of the different conditions. ERPs were implemented during this
task and time-locked to the presentation of each word.

Finally, a short 48-item picture naming task was included. Pictures were black and white
line drawings (Snodgrass & Vanderwart, 1980). Participants were instructed to name the picture
as specifically as possible into the microphone. English monolinguals completed this task prior to
performing the English lexical decision task. L2 learners, on the other hand, completed this task
halfway through the Spanish lexical decision task as a way to induce concurrent language
activation to compare Spanish performance in the first half compared to the second half of the
task.
**Procedure**

Figure 2-1: Schematic of the procedure used in the present study. All participants completed session one in English; only the monolinguals performed the English picture naming task in session one. L2 learners returned for a second session to perform the lexical tasks in Spanish and the English picture naming task.

Figure 2-1 provides a schematic of the order of tasks and which participants completed each. The tasks and materials were not counterbalanced across sessions: all participants performed tasks in English in the first session in order to measure L1 performance in isolation, without the immediate effects of recent L2 use. Materials for the lexical decision task were
matched across languages, but not counterbalanced. Categories for the verbal fluency tasks were the same across languages.

In the first session, participants completed the language history questionnaire and individual difference measures prior to completing the critical language processing tasks. After they completed the language history questionnaire, they performed the O-Span task, AX-CPT, and Flanker. The experimenter described each task to the participant and monitored the practice trials. After the individual difference measures were completed, participants performed the lexical processing tasks in English: English verbal fluency and English lexical decision.

For the verbal fluency task, participants were seated in front of a microphone for recording their responses. They completed a practice trial, colors, followed by the five experimental categories, which were presented in random order. After the semantic fluency task, the final task in the first session for the L2 learners was the English lexical decision task. The English monolinguals performed the English picture-naming task prior to the English lexical decision task. For the lexical decision task, participants were fitted with an ERP cap so that ERP measurements could be time-locked to each item. The ERP system uses a Neuroscan SynAmps2 amplifier with 24-bit analog to digital conversion (Compumedics NeuroScan, Inc., El Paso, TX). This system is an integrated array of low noise amplifiers designed to record high-density EEGs of human brain electrical activity. The cap has 32 electrodes distributed across the scalp. The participant was situated in a sound-attenuated, electrically-shielded booth with a computer, keyboard, button box and microphone. The researchers were in an adjacent room with two computers: one for continuous display of the EEG recordings and impedances and one for stimuli display from the point of view of the participant. After getting situated with the ERP system in place, participants completed the lexical decision task.

Only L2 learners returned for the second session, in which the same lexical processing tasks were administered, but in Spanish. The order of tasks remained the same: Spanish verbal
fluency followed by the Spanish lexical decision. Halfway through the Spanish lexical decision task, participants paused and performed the English picture-naming task. EEG recordings were recorded for the lexical decision task but not for the intervening English picture-naming task.
Chapter 3

Results

As outlined in the introduction, the analyses included group comparisons with the monolingual participants for the tasks performed in English and comparisons with only the two groups of L2 learners for tasks performed in Spanish. A subset of participants were matched on demographic and cognitive measures and then compared on the same dependent measures. Finally, continuous statistics were used to evaluate the relationship between L2 proficiency, L2 sensitivity during L1 processing, and L1 costs within the group of L2 learners.

Data Cleaning and Coding Procedures

For the language history questionnaire, participants were regarded as advanced L2 learners if their average self-rating for Spanish proficiency (across reading, spelling, writing, comprehension, and speaking) was greater than 5 (n = 25). Participants who were actively learning Spanish but self-rated their proficiency as less than 5 were regarded as beginning L2 learners (n = 19). Participants who reported little to no knowledge of an L2, self-rated their Spanish proficiency as less than 5, and considered themselves monolinguals were classified as functionally monolingual (n = 25). Participants who had significant knowledge of a language other than English or Spanish, and participants who reported learning Spanish or any other language in a home setting were excluded (n = 3, already excluded from reported n above).
In the other tasks, data were excluded as outliers if they deviated more than 2.5 standard deviations from the mean. For the O-Span task, one participant’s data were excluded since (s)he did not answer the math problems during the task, resulting in an O-Span score of 0.

In the Flanker task, trials were excluded from contributing to the Flanker Effect if they were more or less than 2.5 standard deviations from the participant’s mean reaction time. The Flanker Effect was calculated by subtracting the (cleaned) mean reaction time for congruent trials from the (cleaned) mean reaction time for incongruent trials. Across participants, one person’s data were excluded whose Flanker Effect was more than 2.5 standard deviations beyond others.

In the AX-CPT task, the primary dependent variable was the behavior shift index (BSI). Higher BSI values indicated a higher reliance on proactive inhibitory control, whereas lower BSI values indicate a higher reliance on reactive inhibitory control. The formula to calculate the BSI is:

$$\frac{(AY \text{ trials RT} - BX \text{ trials RT})}{(AY \text{ trials RT} + BX \text{ trials RT})} + \frac{(AY \text{ trials error rate} - BX \text{ trials error rate})}{(AY \text{ trials error rate} + BX \text{ trials error rate})}$$

No participant’s data were more or less than 2.5 standard deviations from the mean BSI, therefore all participants’ BSI data were included for continuous statistics.

For the lexical decision tasks, the two dependent variables of interest were d’ and reaction time. The d’ score is calculated by subtracting the false alarm rate from the hit rate. The false alarm rate is the number of trials on which a participant incorrectly indicated that a nonword was a word (# false alarms) divided by the total number of nonword trials. The hit rate is the number of trials that a participant correctly identified words (# hits) divided by the total number of word trials. For the English lexical decision task, participants were approaching ceiling values for d’, therefore the measure was transformed to better approximate a normal distribution. Hereafter, statistical analyses referring to d’ as a dependent variable are using the transformed d’.
For the Spanish lexical decision task, the d’ measure satisfied normality without transformation due to the larger variability in d’ scores.

Reaction times in the lexical decision task were cleaned at the trial level to exclude absolute outliers (RTs less than 200 ms or greater than 2000 ms) and relative outliers. Relative outliers were calculated in relation to each participant’s mean reaction time, such that trials were labeled as relative outliers and excluded if they were more or less than 2.5 standard deviations from the participant’s mean RT. In the English lexical decision task, less than 0.01% of the trials were excluded as absolute outliers, 2.7% were excluded as relative outliers, and 2.7% of the trials were incorrectly answered and therefore not included in mean reaction times. English reaction times were log transformed to better approximate a normal distribution. In the Spanish lexical decision task, the same criteria applied, resulting in 0.73% of trials excluded as absolute outliers, 2.4% excluded as relative outliers, and 21% were incorrectly answered and therefore excluded from the mean reaction times. Spanish reaction times were log transformed to better approximate a normal distribution.

Coding procedures for the English and Spanish verbal fluency tasks were described in the Methods section. No data were excluded for the verbal fluency task in either language.

Pre-processing of the ERP waveforms offline involved a 30Hz low-pass filter applied to the continuous EEG record prior to creating averaged epochs of 200 ms prior to the stimulus onset until 800 ms after. The data were then re-referenced to the average of both mastoids and matched on the 200 ms pre-stimulus baseline. Trials contaminated with eye blinks or saccades were excluded from final analyses. In the English lexical decision task, on average 14% of the word trials, 11% of nonword trials, 12.4% of cognate trials, 13.8% of noncognate trials, 15.4% of homograph trials, and 14.7% of the homograph control trials were excluded due to artifact. In the Spanish lexical decision task, on average 7.9% of word trials, 7.6% of nonword trials, 7.6% of cognate trials, 7.1% of noncognate trials, 8.5% of homograph trials, and 8.3% of homograph
control trials were excluded due to artifact. To identify artifacts, any deviations of more than 100 mv in the eye electrode channels were rejected and I additionally viewed the rejected artifacts manually. For the English lexical decision task, only trials that participants correctly answered were included. For the Spanish lexical decision task, ERP waveforms including correct trials only and waveforms including the incorrect trials were created. Both showed the same pattern across groups and conditions. Since the beginning learners had significantly lower accuracy on the Spanish items than the advanced learners, the waveforms including correct trials only were rather unbalanced and messy, therefore the ERPs used for the Spanish lexical decision analyses and plots here were those that included all trials, incorrect trials in addition to correct trials.

The component of interest in the ERP data was the N400, typically found between 300-500ms post-stimulus onset (Duncan et al., 2009). Therefore, the mean amplitude of each participant’s ERPs between 300-500ms were extracted for statistical analyses. In the ANOVAs, four regions of interest (ROIs) were created for the lateral electrodes: left anterior (F7, F3, FT7, FC3), right anterior (F8, F4, FT8, FC4), left posterior (TP7, CP3, P7, P3), and right posterior (TP8, CP4, P8, P4). For the inclusion of ERPs in the continuous statistics of individual differences, the difference waves between two conditions of interest (e.g., difference wave between English words – English nonwords) were averaged over centro-parietal electrodes (C3, Cz, C4, CP3, CP4, P3, Pz, P4).

**Group Comparisons**

**English Lexical Decision Task**

A one-way ANOVA with group (advanced learners-LA, beginning learners-LB, monolinguals-M) as a between-subjects variable and English d’ as the dependent variable
revealed a significant difference between at least two groups, \( F(2, 63) = 3.72, p = .03 \). Post-hoc tests using a Bonferroni correction showed that the advanced learners (\( M = .67, SD = .21 \)) were significantly more accurate than the beginning learners (\( M = .45, SD = .3 \)), but neither group differed from the monolinguals (\( M = .51, SD = .29 \)).

Figure 3-1: Accuracy results (d’) for the three groups in the English lexical decision task. Advanced learners were significantly more accurate the beginning learners and monolinguals, with no difference between the latter two groups.

A one-way ANOVA with group (LA, LB, M) as a between-subjects factor and English reaction times as the dependent variable revealed no difference between any of the groups, \( F(2, 64) = 0.23, p = .8 \). Advanced learners (\( M = 6.46, SD = 0.11 \)) were as fast as beginning learners (\( M = 6.48, SD = 0.16 \)) and monolinguals (\( M = 6.48, SD = 0.17 \)).
A repeated measures ANOVA on reaction times was conducted with word status (word, nonword) as a within-subject factor and group (LA, LB, M) as a between-subject factor. Words (M = 6.41, SD = 0.14) elicited faster reaction times than nonwords (M = 6.53, SD = 0.16), F(1, 64) = 129.4, p < .001. There was no effect of group, F(2, 64) = 0.24, p = .79, nor was there an interaction between group and word status, F(2, 64) = 0.38, p = .7.
A repeated measures ANOVA on accuracy was conducted with word status (word, nonword) as a within-subject factor and group (LA, LB, M) as a between-subject factor. There was a marginal effect of word status, $F(1, 64) = 3.15, p = 0.08$, revealing that participants were slightly more accurate for words ($M = .95, SD = .04$) compared to nonwords ($M = .94, SD = .05$). There was no effect of group, $F(2, 64) = 2.48, p = .09$, nor was there an interaction between group and word status, $F(2, 64) = 0.51, p = .61$.

Figure 3-3: Reaction time differences for words (red) and nonwords (blue) for the three groups in the English lexical decision task. All groups were faster for words compared to nonwords, but did not differ from each other in reaction times.
Two repeated measures ANOVA were conducted on the mean amplitude of the ERPs between 300-500ms post-stimulus, one for the midline electrodes (Fz, Cz, Pz) and one for the ROIs (left anterior, left posterior, right anterior, right posterior).

The midline ANOVA included electrode (Fz, Cz, Pz) and word status (word, nonword) as within-subject factors, and group (LA, LB, M) as a between-subject factor.

Table 3-1: Results of the 3 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for words and nonwords in the English lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
<th>Greenhouse Geisser Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrode</td>
<td>2</td>
<td>114</td>
<td>51.85</td>
<td>&lt; .001*</td>
<td></td>
</tr>
<tr>
<td>Word Status</td>
<td>1</td>
<td>57</td>
<td>99.89</td>
<td>&lt; .001</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>2</td>
<td>57</td>
<td>.65</td>
<td>.53</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-4: Accuracy results for words (red) and nonwords (blue) for the three groups in the English lexical decision task. Despite overall high accuracy rates, participants were marginally more accurate for words than nonwords, but none of the groups differed in accuracy.
From a visual inspection of the mean amplitudes by group, the main effect of word status becomes obvious, revealing that words are, as expected, less negative than nonwords, confirming the presence of an N400 effect in all three groups of participants. Although the group x word status interaction was no longer significant after a Greenhouse-Geisser correction, the interaction is noticeable in the graph presented here, revealing what appears to be a larger difference between words and nonwords for two groups of L2 learners than the monolinguals.

The ROI ANOVA included anteriority (anterior, posterior), hemisphere (left, right), and word status (word, nonword) as within-subject factors, and group (LA, LB, M) as between-subject factors. To evaluate the group x word status interaction, a separate ANOVA was
conducted on the difference wave between words and nonwords to evaluate how the magnitude of the word effect differed across groups, which revealed a main effect of group, $F(2, 57) = 4.07, p = .02$. Paired t-tests using a Bonferroni correction for multiple comparisons showed that all three groups significantly differed from each other in the magnitude of the word effect. Advanced learners had a larger word effect than the beginning learners ($p < .001$) and the monolinguals ($p < .001$), and the beginning learners had a larger word effect than the monolinguals ($p = .01$).

Table 3-2: Results of the $2 \times 2 \times 2 \times 3$ ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for words and nonwords in the English lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>57</td>
<td>0.41</td>
<td>.07</td>
</tr>
<tr>
<td>Word Status</td>
<td>1</td>
<td>57</td>
<td>130.65</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1</td>
<td>57</td>
<td>22.51</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Anteriority</td>
<td>1</td>
<td>57</td>
<td>13.58</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Group x Word Status</td>
<td>2</td>
<td>57</td>
<td>4.07</td>
<td>.02</td>
</tr>
<tr>
<td>Group x Hemisphere</td>
<td>2</td>
<td>57</td>
<td>1.33</td>
<td>.27</td>
</tr>
<tr>
<td>Group x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.09</td>
<td>.92</td>
</tr>
<tr>
<td>Word Status x Hemisphere</td>
<td>1</td>
<td>57</td>
<td>36.37</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Word Status x Anteriority</td>
<td>1</td>
<td>57</td>
<td>0.54</td>
<td>.47</td>
</tr>
<tr>
<td>Hemisphere x Anteriority</td>
<td>1</td>
<td>57</td>
<td>31.48</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Group x Word Status x Hemisphere</td>
<td>2</td>
<td>57</td>
<td>0.08</td>
<td>.92</td>
</tr>
<tr>
<td>Group x Word Status x Anteriority</td>
<td>2</td>
<td>57</td>
<td>1.76</td>
<td>.18</td>
</tr>
<tr>
<td>Group x Hemisphere x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.45</td>
<td>.64</td>
</tr>
<tr>
<td>Word Status x Hemisphere x Anteriority</td>
<td>1</td>
<td>57</td>
<td>16.78</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>
Here the main effect of word status once again is apparent, with words consistently less negative than the nonwords. The effect of anteriority is also noticeable, with posterior electrode sites consistently more positive than the anterior sites.

Figure 3-6: Mean amplitudes (in mv) from 300-500 ms for English words (blue) and nonwords (red) across the four regions of interest for all three groups. Advanced learners had the largest difference between words and nonwords, followed by beginning learners, and then monolinguals.
A visual inspection of the ERP waveforms and scalp distribution confirms the results of the statistical analyses. Plotted below is a representative electrode (Cz) for the three different groups of learners in addition to showing the scalp distribution of the effect of word status.

Advanced learners:

![Graph showing ERP waveforms and scalp distribution for advanced learners.]

Beginning learners:

![Graph showing ERP waveforms and scalp distribution for beginning learners.]

*Student Version of MATLAB*
Monolinguals:

![Waveform and scalp distribution graphs for Monolingual English Words and Monolingual English Nonwords.](Student Version of MATLAB)

Figure 3-7: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -4 to 4 mv) for all three groups showing the difference between English words and nonwords.

Overall, it appears that the behavioral and ERP data concur for the effect of word status. While the ERP data reveal a potential group interaction with word status, the behavioral data show almost the exact same size of a word effect across groups. However, the monolinguals also had overall greater variance in their reaction times for words and nonwords, thus it is reasonable to expect that this behavioral variability is related to the slightly smaller N400 in the group monolingual average on the ERP record which may include jittered processing. All three groups showed strikingly similar scalp distributions for the effect, but again, the advanced learners’ effect seemed to be the strongest, earliest, and most focal.

**Cognates vs. Noncognates**

A repeated measures ANOVA on reaction times was conducted with cognate status (cognate, noncognate) as a within-subject factor and group (LA, LB, M) as a between-subject factor. Cognate status did not significantly impact reaction times, $F(1, 64) = 0.06$, $p = .81$, nor did
group, $F(2, 64) = 0.41, p = .67$. There was not an interaction between groups and cognate status, $F(2, 64) = 0.85, p = .43$.

A repeated measures ANOVA on accuracy was conducted with cognate status (cognate, noncognate) as a within-subject factor and group (LA, LB, M) as a between-subject factor. Participants were more accurate in responding to cognates (M = .96, SD = .05) than to noncognates (M = .95, SD = .05), $F(1, 64) = 12.09, p < .01$. There were no differences between groups in accuracy for cognates and noncognates, $F(2, 64) = 0.25, p = .78$. There was not an interaction between groups and cognate status, $F(2, 64) = 0.64, p = .53$.

Two repeated measures ANOVA were conducted on the mean amplitude of the ERPs between 300-500ms post-stimulus, one for the midline electrodes (Fz, Cz, Pz) and one for the ROIs (left anterior, left posterior, right anterior, right posterior).

The midline ANOVA included electrode (Fz, Cz, Pz) and cognate status (cognate, noncognate) as within-subject factors, and group (LA, LB, M) as a between-subject factor.

Figure 3-8: Accuracy rates for cognates (red) and noncognates (blue) for all three groups in the English lexical decision task. Participants were more accurate for cognates than noncognates, but none of the groups differed in overall accuracy rates.
Table 3-3: Results of the 3 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for cognates and noncognates in the English lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
<th>Greenhouse Geisser Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>57</td>
<td>0.64</td>
<td>.53</td>
<td></td>
</tr>
<tr>
<td>Electrode</td>
<td>2</td>
<td>114</td>
<td>47.09</td>
<td>&lt; .001</td>
<td>*</td>
</tr>
<tr>
<td>Cognate Status</td>
<td>1</td>
<td>57</td>
<td>5.36</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode</td>
<td>4</td>
<td>114</td>
<td>0.32</td>
<td>.86</td>
<td></td>
</tr>
<tr>
<td>Group x Cognate Status</td>
<td>2</td>
<td>57</td>
<td>0.96</td>
<td>.39</td>
<td></td>
</tr>
<tr>
<td>Electrode x Cognate Status</td>
<td>2</td>
<td>114</td>
<td>1.01</td>
<td>.37</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode x Cognate Status</td>
<td>4</td>
<td>114</td>
<td>0.73</td>
<td>.57</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-9:  Mean amplitude values (in mv) from 300-500 ms for English cognates (blue) and noncognates (red) averaged across the three midline electrodes (Fz, Cz, Pz) for all three groups. Cognates were less negative than noncognates prior to the Greenhouse-Geisser correction.

This graph of the mean amplitudes for the different groups reveals the effect of cognate status (before the Greenhouse-Geisser correction) and further hints that a group x cognate status interaction may be present. Advanced learners appear to have reduced negativity for cognates and
the same effect is smaller but emerging within the beginning learners. However, the error bars indicate large variance in the effect, effectively washing out any group effects. Reassuringly, the monolingual group appears to show no distinction between cognates and noncognates, thus confirming their monolingual status and the quality of the lexical characteristic matching.

The ROI ANOVA included anteriority (anterior, posterior), hemisphere (left, right), and cognate status (cognate, noncognate) as within-subject factors, and group (LA, LB, M) as between-subject factors.

Table 3-4: Results of the 2 x 2 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for cognates and noncognates in the English lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>57</td>
<td>0.29</td>
<td>.75</td>
</tr>
<tr>
<td>Cognate Status</td>
<td>1</td>
<td>57</td>
<td>3.3</td>
<td>.07</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1</td>
<td>57</td>
<td>4.01</td>
<td>.05</td>
</tr>
<tr>
<td>Anteriority</td>
<td>1</td>
<td>57</td>
<td>11.82</td>
<td>.001</td>
</tr>
<tr>
<td>Group x Cognate Status</td>
<td>2</td>
<td>57</td>
<td>0.89</td>
<td>.41</td>
</tr>
<tr>
<td>Group x Hemisphere</td>
<td>2</td>
<td>57</td>
<td>0.47</td>
<td>.62</td>
</tr>
<tr>
<td>Group x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.14</td>
<td>.87</td>
</tr>
<tr>
<td>Cognate Status x Hemisphere</td>
<td>1</td>
<td>57</td>
<td>0.14</td>
<td>.72</td>
</tr>
<tr>
<td>Cognate Status x Anteriority</td>
<td>1</td>
<td>57</td>
<td>0.12</td>
<td>.73</td>
</tr>
<tr>
<td>Hemisphere x Anteriority</td>
<td>1</td>
<td>57</td>
<td>61.28</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Group x Cognate Status x Hemisphere</td>
<td>2</td>
<td>57</td>
<td>0.31</td>
<td>.74</td>
</tr>
<tr>
<td>Group x Cognate Status x Anteriority</td>
<td>2</td>
<td>57</td>
<td>1.44</td>
<td>.25</td>
</tr>
<tr>
<td>Group x Hemisphere x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.74</td>
<td>.48</td>
</tr>
<tr>
<td>Cognate Status x Hemisphere x Anteriority</td>
<td>1</td>
<td>57</td>
<td>0.15</td>
<td>.7</td>
</tr>
</tbody>
</table>
Again, a visual inspection reveals the advanced learners appear to be sensitive to the presence of cognates in their L1. At lateral sites, the emerging effect in beginning learners is smaller, but the ANOVA indicates that the effect of cognate status is, indeed, a marginal effect.

The ERP waveforms and scalp distributions confirm the statistical analyses. The advanced learners show a clear N400 response for cognates compared to noncognates, as suggested by the statistical analyses. Similarly, the beginning learners show a very small but...
emerging effect, made clearer in the scalp distribution than the waveform itself. The scalp distribution and waveform of the monolinguals reveals no sensitivity to the presence of Spanish cognates.

Advanced learners:

![Graph showing scalp distribution and waveform for advanced learners.]

Beginning learners:

![Graph showing scalp distribution and waveform for beginning learners.]

---

**Advanced Learner English Cognates**

**Advanced Learner English Noncognates**

**Beginning Learner English Cognates**

**Beginning Learner English Noncognates**

---

**Student Version of MATLAB**

---

**CZ**
Monolinguals:

Figure 3-11: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -1 to 1 mv) for all three groups showing differences between English cognates and noncognates.

The ERP and behavioral data do not line up as clearly for the cognate status as they did for word status. There was no behavioral difference between cognates and noncognates in reaction times. There was a significant difference between cognates and noncognates in accuracy measures, but this did not interact with group, meaning that monolinguals were also more accurate for cognates than noncognates. However, the ERPs reveal that the L2 learners, especially the advanced learners, are sensitive to the language ambiguity present in cognate words, as demonstrated through visual inspection of the mean amplitudes, waveforms, and scalp distributions, whereas the monolingual waveforms and scalp distributions are not consistent with an N400 effect. The similarity across groups for the word vs. nonword distinction was expected given that English is the L1 for all the participants, making the distinction between words and nonwords a within-language effect. The difference between L2 learners and monolinguals found in the waveforms and scalp distributions for the cognates vs. noncognates was also predicted, given that the cognate manipulation is a between-language effect to which only the L2 learners should be sensitive. The unexpected results from these analyses are that monolinguals also benefit
in behavioral accuracy from the presence of cognates, and that both groups of learners are demonstrating an emerging N400 effect for cognates vs. noncognates in their L1 processing. For advanced learners, the effect is stronger and more noticeable, but there appears to also be a very small effect in the beginning learners.

**Homographs vs. Controls**

A repeated measures ANOVA on reaction times was conducted with homograph status (homograph, control) as a within-subject factor and group (LA, LB, M) as a between-subjects factor. Homograph status did not significantly impact reaction times, $F(1, 64) = 0.62, p = .43$, nor did group, $F(2, 64) = 0.26, p = .77$. There was not an interaction between groups and homograph status, $F(2, 64) = 0.87, p = .42$.

A repeated measures ANOVA on accuracy was conducted with homograph status (homograph, control) as a within-subject factor and group (LA, LB, M) as a between-subjects factor. Participants were more accurate in responding to homograph controls ($M = .96, SD = .05$) than to homographs ($M = .94, SD = .06$), $F(1, 64) = 8.96, p < .01$. There were no differences between groups in accuracy for homographs and controls, $F(2, 64) = 1.6, p = .21$. There was not an interaction between group and homograph status, $F(2, 64) = 1.01, p = .37$. These results reveal that monolinguals were equally affected by the presence of homographs, indicating that the effect may not reflect language ambiguity but rather lexical properties of these L1 words and the matched control words.
Two repeated measures ANOVA were conducted on the mean amplitude of the ERPs between 300-500ms post-stimulus, one for the midline electrodes (Fz, Cz, Pz) and one for the ROIs (left anterior, left posterior, right anterior, right posterior).

The midline ANOVA included electrode (Fz, Cz, Pz) and homograph status (homograph, control) as within-subject factors, and group (LA, LB, M) as a between-subjects factor.

Figure 3-12: Accuracy rates for homographs (red) and controls (blue) for all three groups in the English lexical decision task. Participants were more accurate for controls than homographs, including monolinguals.

Two repeated measures ANOVA were conducted on the mean amplitude of the ERPs between 300-500ms post-stimulus, one for the midline electrodes (Fz, Cz, Pz) and one for the ROIs (left anterior, left posterior, right anterior, right posterior).

The midline ANOVA included electrode (Fz, Cz, Pz) and homograph status (homograph, control) as within-subject factors, and group (LA, LB, M) as a between-subjects factor.

Table 3-5: Results of the 3 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for homographs and controls in the English lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
<th>Greenhouse Geisser Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>57</td>
<td>1.04</td>
<td>.36</td>
<td></td>
</tr>
<tr>
<td>Electrode</td>
<td>2</td>
<td>114</td>
<td>48.86</td>
<td>&lt; .001*</td>
<td></td>
</tr>
<tr>
<td>Homograph Status</td>
<td>1</td>
<td>57</td>
<td>0.09</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode</td>
<td>4</td>
<td>114</td>
<td>0.21</td>
<td>.94</td>
<td></td>
</tr>
</tbody>
</table>
The homographs appear to be having differential effects in different groups: no effect in advanced learners, facilitation in beginning learners, and interference in monolinguals. The error bars once again reveal the lack of any statistical significance, but the trends in the data are unexpected.

The ROI ANOVA included anteriority (anterior, posterior), hemisphere (left, right), and homograph status (homograph, control) as within-subject factors, and group (LA, LB, M) as between-subject factors.

Figure 3-13: Mean amplitude values (in mv) from 300-500 ms for English homographs (blue) and controls (red) averaged across the three midline electrodes (Fz, Cz, Pz) for all three groups. Homographs show different effects in different groups, though not statistically significant.
Table 3-6: Results of the 2 x 2 x 2 x 3 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for homographs and controls in the English lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>57</td>
<td>1.04</td>
<td>.36</td>
</tr>
<tr>
<td>Homograph Status</td>
<td>1</td>
<td>57</td>
<td>0.02</td>
<td>.89</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1</td>
<td>57</td>
<td>9.9</td>
<td>&lt;.003</td>
</tr>
<tr>
<td>Anteriority</td>
<td>1</td>
<td>57</td>
<td>13.94</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Group x Homograph Status</td>
<td>2</td>
<td>57</td>
<td>1.36</td>
<td>.26</td>
</tr>
<tr>
<td>Group x Hemisphere</td>
<td>2</td>
<td>57</td>
<td>1.72</td>
<td>.19</td>
</tr>
<tr>
<td>Group x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.13</td>
<td>.88</td>
</tr>
<tr>
<td>Homograph Status x Hemisphere</td>
<td>1</td>
<td>57</td>
<td>0.48</td>
<td>.49</td>
</tr>
<tr>
<td>Homograph Status x Anteriority</td>
<td>1</td>
<td>57</td>
<td>2.27</td>
<td>.14</td>
</tr>
<tr>
<td>Hemisphere x Anteriority</td>
<td>1</td>
<td>57</td>
<td>21.78</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Group x Homograph Status x Hemisphere</td>
<td>2</td>
<td>57</td>
<td>0.38</td>
<td>.69</td>
</tr>
<tr>
<td>Group x Homograph Status x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.16</td>
<td>.85</td>
</tr>
<tr>
<td>Group x Hemisphere x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.54</td>
<td>.58</td>
</tr>
<tr>
<td>Homograph Status x Hemisphere x Anteriority</td>
<td>1</td>
<td>57</td>
<td>0.23</td>
<td>.63</td>
</tr>
<tr>
<td>Group x Homograph Status x Hemisphere x Anteriority</td>
<td>2</td>
<td>57</td>
<td>0.43</td>
<td>.65</td>
</tr>
</tbody>
</table>
Again, the statistical analyses reveal that any differences between groups of conditions are not reliable, but similar trends persist across groups. A visual inspection of the waveforms and scalp distribution confirms the statistical analyses of the ERPs and behavioral measures.
Advanced learners:

![Graph for Advanced Learners]

Beginning learners:

![Graph for Beginning Learners]

Monolinguals:

![Graph for Monolinguals]
The scalp distribution for the monolinguals seems to suggest an overall negativity, rather than an effect that resembles the N400, whereas the emerging effect in the beginning learners does appear to be more focused and arises within the central electrodes, as a typical N400 would. An interesting discrepancy arises between the visual ERPs and the behavioral analyses, namely that all groups demonstrated interference for homographs in the form of reduced accuracy for homographs compared to controls, yet the beginning learners show emerging facilitation at the neural level in the form of a very small but reduced N400 for homographs compared to controls. The monolingual ERPs and behavioral data are in line with each other, both suggesting interference of some sort, but these monolinguals should not be sensitive to words that share overlap with Spanish. Together, these data suggest there may have been correlated lexical factors influencing the results for both L2 learners and monolinguals that are unrelated to Spanish knowledge. The differing patterns of activation across the groups, rather than a homogenous effect of the lexical characteristics, may tentatively suggest that the interactions between L1 and L2 knowledge modulated the effects found in the L2 learners, but this conclusion cannot be confirmed with the data presented here.

**Spanish Lexical Decision Task**

A t-test was conducted to compare d’ scores across the two groups of learners on the Spanish lexical decision task, revealing that advanced learners (M = .67, SD = .09) were more accurate in discriminating words and nonwords in Spanish than the beginning learners (M = .46, SD = .18), t(24) = 4.6, p < .001. Another t-test was conducted on reaction times across the two
groups of learners, revealing that advanced learners (M = 6.63, SD = 0.15) are also marginally faster than beginning learners (M = 6.73, SD = 0.19), t(32) = 1.76, p = .08.

**Words vs. Nonwords**

A repeated measures ANOVA was conducted on reaction times with word status (word, nonword) as a within-subject factor and group (LA, LB) as a between-subjects factor. Both groups were faster to respond to words (M = 6.65, SD = 0.18) than to nonwords (M = 6.7, SD = 0.19), F(1, 40) = 5.04, p = .03. There was a marginal effect of group, with advanced learners (M = 6.63, SD = 0.16) responding faster than beginning learners (M = 6.73, SD = 0.2), F(1, 40) = 3.89, p = .055. There was not an interaction between group and word status, F(1, 40) = 0.55, p = .46.

Figure 3-16: Reaction times to words (red) and nonwords (blue) for beginning and advanced learners in the Spanish lexical decision task. Larger variance for beginning learners in response to words suggests that they may be guessing (quickly) for some correct responses and fully processing the words (more slowly) other times.
A repeated measures ANOVA was conducted on accuracy with word status (word, nonword) as a within-subject factor and group (LA, LB) as a between-subjects factor. Both groups were more accurate in responding to words (M = .69, SD = .12) than to nonwords (M = .83, SD = .15), F(1, 40) = 20.84, p < .001. Advanced learners (M = .8, SD = .12) were more accurate than beginning learners (M = .69, SD = .16), F(1, 40) = 23.7, p < .001. There was not an interaction between group and word status, F(1, 40) = 0.66, p = .42.

![Figure 3-17](image.png)

Figure 3-17: Accuracy rates for Spanish words (red) and nonwords (blue) for beginning and advanced learners. Again, the large variance in beginning learners suggests a high rate of guessing.

Two repeated measures ANOVA were conducted on the mean amplitude of the ERPs between 300-500ms post-stimulus, one for the midline electrodes (Fz, Cz, Pz) and one for the ROIs (left anterior, left posterior, right anterior, right posterior).

The midline ANOVA included electrode (Fz, Cz, Pz) and word status (word, nonword) as within-subject factors, and group (LA, LB) as a between-subjects factor.
Table 3-7: Results of the 3 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for words and nonwords in the Spanish lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
<th>Greenhouse Geisser Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>35</td>
<td>0.01</td>
<td>.92</td>
<td></td>
</tr>
<tr>
<td>Electrode</td>
<td>2</td>
<td>70</td>
<td>33.07</td>
<td>&lt; .001</td>
<td>*</td>
</tr>
<tr>
<td>Word Status</td>
<td>1</td>
<td>35</td>
<td>68.33</td>
<td>&lt; .001</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode</td>
<td>2</td>
<td>70</td>
<td>1.71</td>
<td>.19</td>
<td></td>
</tr>
<tr>
<td>Group x Word Status</td>
<td>1</td>
<td>35</td>
<td>2.05</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>Electrode x Word Status</td>
<td>2</td>
<td>70</td>
<td>0.05</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode x Word Status</td>
<td>2</td>
<td>70</td>
<td>2.19</td>
<td>.12</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-18: Mean amplitudes (in mv) from 300-500 ms for Spanish words (blue) and nonwords (red) for beginning and advanced learners across midline electrodes (Fz, Cz, Pz). Both groups showed a main effect of word status but did not differ from each other.
The main effect of word status is clear in the graph; both groups of learners show an N400 effect for Spanish nonwords compared to words. Both groups are proficient enough to perform the task. It is surprising that there is no group interaction, statistically or visually, since we might expect the advanced learners to be more sensitive to the word/nonword distinction in Spanish.

The ROI ANOVA included anteriority (anterior, posterior), hemisphere (left, right), and word status (word, nonword) as within-subject factors, and group (LA, LB) as between-subject factors.

Table 3-8: Results of the 2 x 2 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for words and nonwords in the Spanish lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>35</td>
<td>0.01</td>
<td>.94</td>
</tr>
<tr>
<td>Word Status</td>
<td>1</td>
<td>35</td>
<td>54.85</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1</td>
<td>35</td>
<td>50.92</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Anteriority</td>
<td>1</td>
<td>35</td>
<td>7.01</td>
<td>.01</td>
</tr>
<tr>
<td>Group x Word Status</td>
<td>1</td>
<td>35</td>
<td>1.5</td>
<td>.12</td>
</tr>
<tr>
<td>Group x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>1.57</td>
<td>.22</td>
</tr>
<tr>
<td>Group x Anteriority</td>
<td>1</td>
<td>35</td>
<td>1.38</td>
<td>.25</td>
</tr>
<tr>
<td>Word Status x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>1.2</td>
<td>.28</td>
</tr>
<tr>
<td>Word Status x Anteriority</td>
<td>1</td>
<td>35</td>
<td>8.57</td>
<td>.01</td>
</tr>
<tr>
<td>Hemisphere x Anteriority</td>
<td>1</td>
<td>35</td>
<td>9.01</td>
<td>.01</td>
</tr>
<tr>
<td>Group x Word Status x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>0.93</td>
<td>.34</td>
</tr>
<tr>
<td>Group x Word Status x Anteriority</td>
<td>1</td>
<td>35</td>
<td>0.54</td>
<td>.47</td>
</tr>
</tbody>
</table>
Figure 3-19: Mean amplitudes (in µv) from 300-500 ms for Spanish words (blue) and nonwords (red) across the four regions of interest for advanced and beginning learners. The main effect of word status is apparent, while there was no effect of group or any group interactions.
The main effect of word status is persistent throughout the ERP data, across electrode regions of interest. Again, the difference in magnitude between the two groups remains consistent, and the advanced learners do not appear to show a larger difference than beginning learners.

The visual waveforms and scalp plots similarly suggest that both groups can reliably distinguish words and nonwords in Spanish.

Advanced learners:

![Advanced Learner Waveforms and Scalp Plots](image1)

Beginning learners:

![Beginning Learner Waveforms and Scalp Plots](image2)
The scalp distributions for the effect reveal further differences between the two groups. The advanced learners, as expected, show a focal N400 effect in centro-parietal electrodes, whereas the nonword negativity (or, word positivity) appears in more anterior electrodes in beginning learners and is much less focal. Overall, the behavioral and electrophysiological data agree, both suggesting that Spanish nonwords are more difficult than Spanish words, and that advanced learners are overall faster, more accurate, and slightly more sensitive to the distinction between words and nonwords, though not statistically.

**Cognates vs. Noncognates**

A repeated measures ANOVA was conducted on reaction times with cognate status (cognate, noncognate) as a within-subject factor and group (LA, LB) as a between-subjects factor. Participants were faster to respond to cognates (M = 6.64, SD = 0.2) than to noncognates (M = 6.68, SD = 0.19), F(1, 40) = 10.03, p < .01. There is a marginal effect of group, with advanced learners (M = 6.62, SD = 0.15) responding faster than beginning learners (M = 6.72, SD = 0.23), F(1, 40) = 3.21, p = .08. There was not an interaction between group and cognate status, F(1, 40) = 0.02, p = .9.

---

**Figure 3-20**: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -2 to 2 mv for beginning learners, -3 to 3 for advanced learners) for both groups showing differences between Spanish words and nonwords.
A repeated measures ANOVA was conducted on accuracy with cognate status (cognate, noncognate) as a within-subject factor and group (LA, LB) as a between-subjects factor. Both groups were more accurate in responding to cognates ($M = .78$, $SD = .18$) than to noncognates ($M = .68$, $SD = .14$), $F(1, 40) = 24.12$, $p < .001$. Advanced learners ($M = .79$, $SD = .13$) were more accurate than beginning learners ($M = .66$, $SD = .19$), $F(1, 40) = 10.21$, $p < .01$. There was an interaction between group and word status, $F(1, 40) = 4.14$, $p = .049$. Pairwise t-tests between word types within each group revealed that advanced learners were significantly more accurate for cognates compared to noncognates, $t(44) = 4.14$, $p < .001$, whereas the beginning learners were not significantly more accurate for cognates compared to noncognates, $t(30) = .84$, $p = .41$. This lack of a cognate effect in beginning learners is unexpected given that learners typically

Figure 3-21: Reaction time results for Spanish cognates (red) and noncognates (blue) for advanced and beginning learners. Cognates elicited faster reaction times than noncognates, and advanced learners responded marginally faster than beginning learners. No cognate facilitation is present in the beginning learners, likely due to the large variance and the low accuracy rate, which suggests they were guessing, thereby washing out any effects.
show larger cognate effects earlier in L2 learning (Tonzar, Lotto & Job, 2009), however the low accuracy rates for words and nonwords suggests a high rate of guessing. Guesses are often faster responses than fully processing and making a lexical decision about an item. The mix of guesses and L2 word processing in the beginning learners created large variability in their reaction time and accuracy rates, which most likely washed out any cognate effect.

Two repeated measures ANOVA were conducted on the mean amplitude of the ERPs between 300-500ms post-stimulus, one for the midline electrodes (Fz, Cz, Pz) and one for the ROIs (left anterior, left posterior, right anterior, right posterior).

The midline ANOVA included electrode (Fz, Cz, Pz) and cognate status (cognate, noncognate) as within-subject factors, and group (LA, LB) as a between-subjects factor.

Figure 3-22: Accuracy rates for Spanish cognates (red) and noncognates (blue) for advanced and beginning learners. Participants were more accurate for cognates than noncognates. Advanced learners were more accurate than beginning learners. Though the magnitude of accuracy differences across groups looks comparable, the variance in the beginning group rendered the difference between cognates and noncognates insignificant for that group.
Table 3-9: Results of the 3 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for cognates and noncognates in the Spanish lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
<th>Greenhouse Geisser Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>35</td>
<td>0.35</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>Electrode</td>
<td>2</td>
<td>70</td>
<td>28.36</td>
<td>&lt; .001</td>
<td>*</td>
</tr>
<tr>
<td>Cognate Status</td>
<td>1</td>
<td>35</td>
<td>5.73</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode</td>
<td>2</td>
<td>70</td>
<td>1.49</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>Group x Cognate Status</td>
<td>1</td>
<td>35</td>
<td>4.09</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>Electrode x Cognate Status</td>
<td>2</td>
<td>70</td>
<td>2.09</td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode x Cognate Status</td>
<td>2</td>
<td>70</td>
<td>0.22</td>
<td>.8</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-23: Mean amplitudes (in mv) from 300-500 ms for Spanish cognates (blue) and noncognates (red) for beginning and advanced learners across midline electrodes (Fz, Cz, Pz). The main effect of cognate status is driven by the cognate effect present in the advanced learners.
Interestingly, the beginning learners do not appear to be sensitive to cognates whereas the advanced learners are showing a clear cognate effect. This is also reflected in the group x cognate status interaction, which did not survive the Greenhouse-Geisser correction yet is visually present all the same.

The ROI ANOVA included anteriority (anterior, posterior), hemisphere (left, right), and cognate status (cognate, noncognate) as within-subject factors, and group (LA, LB) as between-subject factors.

Table 3-10: Results of the 2 x 2 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for cognates and noncognates in the Spanish lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>35</td>
<td>0.34</td>
<td>.56</td>
</tr>
<tr>
<td>Cognate Status</td>
<td>1</td>
<td>35</td>
<td>4.54</td>
<td>.04</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1</td>
<td>35</td>
<td>42.02</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Anteriority</td>
<td>1</td>
<td>35</td>
<td>4.66</td>
<td>.04</td>
</tr>
<tr>
<td>Group x Cognate Status</td>
<td>1</td>
<td>35</td>
<td>1.93</td>
<td>.17</td>
</tr>
<tr>
<td>Group x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>1.6</td>
<td>.21</td>
</tr>
<tr>
<td>Group x Anteriority</td>
<td>1</td>
<td>35</td>
<td>0.87</td>
<td>.36</td>
</tr>
<tr>
<td>Cognate Status x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>0.39</td>
<td>.53</td>
</tr>
<tr>
<td>Cognate Status x Anteriority</td>
<td>1</td>
<td>35</td>
<td>4.85</td>
<td>.03</td>
</tr>
<tr>
<td>Hemisphere x Anteriority</td>
<td>1</td>
<td>35</td>
<td>11.2</td>
<td>.002</td>
</tr>
<tr>
<td>Group x Cognate Status x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>1.11</td>
<td>.3</td>
</tr>
<tr>
<td>Group x Cognate Status x Anteriority</td>
<td>1</td>
<td>35</td>
<td>0.03</td>
<td>.86</td>
</tr>
<tr>
<td>Group x Hemisphere x Anteriority</td>
<td>1</td>
<td>35</td>
<td>0.37</td>
<td>.55</td>
</tr>
</tbody>
</table>
The cognate status x anteriority interaction appears in this figure, revealing that the magnitude of the cognate effect is actually larger at anterior electrode sites than posterior electrode sites, which is unusual. Moreover, there is no group effect or interaction, despite the
fact that the advanced learners seem to be more sensitive to the cognates both in the midline electrodes and behaviorally.

A visual inspection of the waveforms and scalp distribution displays a similar pattern to the behavioral and midline data, specifically that advanced learners may capitalize on the presence of cognates more so than beginning learners.

Advanced learners:

![Waveforms and scalp distributions for advanced learners showing differences between Spanish cognates and noncognates.]

Beginning learners:

![Waveforms and scalp distributions for beginning learners showing differences between Spanish cognates and noncognates.]

Figure 3-25: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -1 to 1 mv for beginning learners, -2 to 2 for advanced learners) for both groups showing differences between Spanish cognates and noncognates.
These data from the Spanish cognates is particularly interesting given the role of the L1 according to the RHM (Kroll & Stewart, 1994), which proposes that L2 words are translated via the L1 to access the concept at lower proficiency. When the L2 and L1 word are identical or share significant overlap, as in the case of cognates, both reaction time and neural processing should be facilitated. The scalp plot of the advanced learners here shows that in the earlier time window (300-400ms), the cognate effect may be slightly stronger but also appears in anterior regions, whereas in a slightly later time window (400-500ms), the effect is present at the predicted centro-parietal sites, but is also weaker. The interpretation of the behavioral analyses of the beginning learners’ performance was that the large variance in performance due to guessing eliminated any sensitivity to a cognate effect. However, the beginning learners’ waveform and scalp distribution confirm the behavioral analyses in demonstrating that beginning learners are not facilitated by the presence of English (L1) cognates during Spanish (L2) processing. Only the midline ERP analyses revealed the group x cognate status interaction that visually appears throughout the behavioral and ERP data.

**Homographs vs. Controls**

Before describing the analyses conducted on the Spanish homograph data, it is worth noting that the lexical decision materials were not counterbalanced across languages. Therefore, the item variability found within the set of English homographs that suggested lexical level effects is not necessarily true within the Spanish homographs and matched control words.

A repeated measures ANOVA was conducted on reaction times with homograph status (homograph, control) as a within-subject factor and group (LA, LB) as a between-subjects factor. There was no effect of homograph status on participant response times, $F(1, 40) = 0.46, p = .5$. 

There was also no effect of group, $F(1, 40) = 1.5, p = .23$. There was not an interaction between group and homograph status, $F(1, 40) = 1.81, p = .19$.

A repeated measures ANOVA was conducted on accuracy with homograph status (homograph, control) as a within-subject factor and group (LA, LB) as a between-subject factor. Both groups were more accurate in responding to homographs ($M = .72, SD = .12$) than to homograph controls ($M = .57, SD = .11$), $F(1, 40) = 121.2, p < .001$. Advanced learners ($M = .66, SD = .13$) were no more accurate than beginning learners ($M = .62, SD = .15$), $F(1, 40) = 1.36, p = .25$. There was an interaction between group and homograph status, $F(1, 40) = 7.05, p = .01$.

Pairwise t-tests revealed that advanced learners were significantly more accurate for homographs compared to control words, $t(46) = 7.26, p < .001$, and beginning learners were also significantly more accurate for homographs compared to control words, $t(34) = 2.4, p = .022$. However, the difference in accuracy between homographs and control words was significantly larger in advanced learners than beginning learners, $t(38) = 2.68, p = .01$.

Figure 3-26: Accuracy rates for Spanish homographs (red) and controls (blue) for advanced and beginning learners. Both groups have higher accuracy rates for homographs compared to controls, but once again, the large variance within the beginning learners’ responses reduced the statistical difference between the two word types.
Two repeated measures ANOVA were conducted on the mean amplitude of the ERPs between 300-500 ms post-stimulus, one for the midline electrodes (Fz, Cz, Pz) and one for the ROIs (left anterior, left posterior, right anterior, right posterior).

The midline ANOVA included electrode (Fz, Cz, Pz) and homograph status (homograph, control) as within-subject factors, and group (LA, LB) as a between-subject factor. A simple effects analysis was conducted on the group x electrode x homograph status interaction. Within the advanced learners, only the electrode remained significant. Within the beginning learners, an electrode x homograph status interaction revealed that homographs elicited less negativity than control words at electrode Fz, but more negativity than control words in electrode Cz, with the largest negativity in electrode Pz.

Table 3-11: Results of the 3 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over midline electrodes for homographs and controls in the Spanish lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
<th>Greenhouse Geisser Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>35</td>
<td>0.02</td>
<td>.88</td>
<td></td>
</tr>
<tr>
<td>Electrode</td>
<td>2</td>
<td>70</td>
<td>28.29</td>
<td>&lt; .001</td>
<td>*</td>
</tr>
<tr>
<td>Homograph Status</td>
<td>1</td>
<td>35</td>
<td>0.2</td>
<td>.65</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode</td>
<td>2</td>
<td>70</td>
<td>2.51</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>Group x Homograph Status</td>
<td>1</td>
<td>35</td>
<td>0.95</td>
<td>.34</td>
<td></td>
</tr>
<tr>
<td>Electrode x Homograph Status</td>
<td>2</td>
<td>70</td>
<td>2.2</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>Group x Electrode x Homograph Status</td>
<td>2</td>
<td>70</td>
<td>4.36</td>
<td>.02</td>
<td>*</td>
</tr>
</tbody>
</table>
The Spanish homograph data reveal an interaction between homograph status, electrode, and group, which can be seen in Figure 3-27. At midline electrode sites, homographs produce interference across electrodes in advanced learners, since homographs are more negative than the control words, whereas follow-up analyses revealed that beginning learners demonstrate facilitation (less negativity for homographs) in anterior electrodes and interferences (greater negativity for homographs) in posterior electrodes. The group x homograph status interaction was
not reliable due to the high level of variability in neural responses to homographs, but the patterns are evident.

The ROI ANOVA included anteriority (anterior, posterior), hemisphere (left, right), and homograph status (homograph, control) as within-subject factors, and group (LA, LB) as between-subject factors.

Table 3-12: Results of the 2 x 2 x 2 x 2 ANOVA conducted on the mean amplitude of the ERPs between 300-500ms over ROI electrodes for homographs and controls in the Spanish lexical decision task.

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1</td>
<td>35</td>
<td>0.03</td>
<td>.86</td>
</tr>
<tr>
<td>Homograph Status</td>
<td>1</td>
<td>35</td>
<td>0.33</td>
<td>.57</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1</td>
<td>35</td>
<td>54.95</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Anteriority</td>
<td>1</td>
<td>35</td>
<td>4.5</td>
<td>.04</td>
</tr>
<tr>
<td>Group x Homograph Status</td>
<td>1</td>
<td>35</td>
<td>0.44</td>
<td>.51</td>
</tr>
<tr>
<td>Group x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>2.09</td>
<td>.16</td>
</tr>
<tr>
<td>Group x Anteriority</td>
<td>1</td>
<td>35</td>
<td>2.07</td>
<td>.16</td>
</tr>
<tr>
<td>Homograph Status x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>0.07</td>
<td>.79</td>
</tr>
<tr>
<td>Homograph Status x Anteriority</td>
<td>1</td>
<td>35</td>
<td>1.96</td>
<td>.17</td>
</tr>
<tr>
<td>Hemisphere x Anteriority</td>
<td>1</td>
<td>35</td>
<td>5.26</td>
<td>.03</td>
</tr>
<tr>
<td>Group x Homograph Status x Hemisphere</td>
<td>1</td>
<td>35</td>
<td>0.01</td>
<td>.92</td>
</tr>
<tr>
<td>Group x Homograph Status x Anteriority</td>
<td>1</td>
<td>35</td>
<td>3.49</td>
<td>.07</td>
</tr>
<tr>
<td>Group x Hemisphere x Anteriority</td>
<td>1</td>
<td>35</td>
<td>0.01</td>
<td>.92</td>
</tr>
<tr>
<td>Homograph Status x Hemisphere x Anteriority</td>
<td>1</td>
<td>35</td>
<td>0.01</td>
<td>.92</td>
</tr>
</tbody>
</table>
Figure 3-28: Mean amplitudes (in mv) from 300-500 ms for Spanish homographs (blue) and controls (red) across the four regions of interest for advanced and beginning learners. As found in the midline analysis, beginning learners show reduced negativity for homographs in posterior regions and increased negativity in anterior regions. Advanced learners show homograph interference (in the form of increased negativity) across the scalp.
The marginal group x homograph status x anteriority interaction can be seen in this figure, which shows a particularly interesting pattern of results for the beginning learners. In the midline analyses, it looked like beginning and advanced learners differed in whether they experienced interference or facilitation. The ROI analyses reveal that beginning learners display greater negativity for homographs compared to controls (i.e. interference) in anterior electrode sites, but the opposite pattern (i.e. facilitation) at posterior sites. Advanced learners, on the other hand, show greater negativity for homographs than matched control words at all sites.

A visual inspection of the waveforms and scalp distributions elucidates these patterns.

Advanced learners:

Beginning learners:

Figure 3-29: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-
Looking specifically to centro-parietal electrodes, the opposing pattern between advanced and beginning learners is apparent. The waveforms show small differences in either direction, which help explain why such small effects are undetected by statistical analyses of homograph status. The behavioral data and ERP data are not necessarily in agreement, since advanced learners were, in fact, more accurate in responding to homographs compared to matched control words, and not at the cost of speed since there were no differences in reaction times based on homograph status or group.

**Verbal Fluency**

For English verbal fluency, a one-way ANOVA was conducted with group (LA, LB, M) as a between-subject factor. The analysis revealed no differences between groups in the number of exemplars produced, $F(2, 54) = 0.33$, $p = .72$.

For Spanish verbal fluency, a t-test comparing advanced and beginning learners revealed that advanced learners ($M = 8.17$, $SD = 3.07$) produced more Spanish exemplars than the beginning learners ($M = 4.96$, $SD = 4$), $t(27) = 2.71$, $p = .01$. 

500 ms (scale: -1 to 1) for both groups showing differences between Spanish homographs and controls.
English Picture Naming

To compare response times for English picture naming, a one-way ANOVA was conducted with group (LA, LB, M) as a between-subject factor. There were no differences between group response times to naming English pictures, $F(2, 52) = 0.4$, $p = .67$.

Individual Difference Measures

A one-way ANOVA conducted on O-Span scores revealed no differences between groups, $F(2, 65) = 1.84$, $p = .17$. Similarly, there were no group differences in the magnitude of the Flanker effect, $F(2, 63) = 1$, $p = .37$, or in the reliance on reactive vs. proactive inhibitory control, as measured by BSI from the AX-CPT task, $F(2, 60) = .24$, $p = .78$.

Figure 3-30: Verbal fluency performance in English (left) and Spanish (right). No groups differed in performance in the English verbal fluency, whereas advanced learners performed better than beginning learners in Spanish.
Matched Subgroup Analysis

In order to compare across groups with more confidence that differences found between groups are due to language knowledge, exposure, and experience rather than other demographic or cognitive factors, a subset of beginning learners and advanced learners were matched with monolinguals. This created a fully crossed design, with some monolinguals (MA) matched to advanced learners (LA) and some monolinguals (MB) matched to beginning learners (LB). With 11 participants per group, there is not much power to detect subtle differences between groups, but since they are matched, paired analyses increase the power to detect these differences.

English Lexical Decision Task

To analyze group differences in English accuracy, a 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) ANOVA was conducted on the transformed English d’ measure. There were no differences between the advanced groups (M = .58, SD = .27) or the beginning groups (M = .59, SD = .26), F(1, 39) = 0.01, p = .92. There were also no differences between the L2 learners (M = .62, SD = .24) and monolinguals (M = .54, SD = .28), F(1, 39) = 0.94, p = .34. There was, however, a significant interaction between language status and level, F(1, 39) = 4.23, p = .046. A paired t-test revealed that advanced L2 learners (M = .69, SD = .23) were more accurate than the matched monolinguals (M = .46, SD = .26), t(10) = 2.05, p = .07, but beginning learners were no different than the matched monolinguals, t(9) = 0.62, p = .55.
To analyze group differences in reaction times, a 2 (advanced, beginning) x 2 (learner, monolingual) ANOVA was conducted. No differences were found between advanced (M = 6.49, SD = 0.16) and beginning (M = 6.42, SD = 0.12) groups, F(1, 39) = 2.85, p = .1. There were also no differences between learners (M = 6.44, SD = 0.1) and monolinguals (M = 6.48, SD = 0.18), F(1,39) = 0.56, p = .46. The interaction between language status and level did not reach significance, F(1, 39) = 0.45, p = .51.

**Words vs. Nonwords**

A 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) x 2 (word status: word, nonword) ANOVA was conducted on reaction times with level and language status as between-subject variables and word status as a within-subject variable. A main effect of word status revealed that words (M = 6.4, SD = 0.14) elicited faster reaction times than nonwords (M =
6.51, SD = 0.15), F(1, 39) = 77.59, p < .001. There was no effect of level, F(1, 39) = 2.83, p = .1, indicating that advanced learners and matched monolinguals (M = 6.49, SD = 0.17) responded no faster or slower than beginning learners and matched monolinguals (M = 6.42, SD = 0.13). There was also no main effect of language status, F(1, 39) = 0.56, p = .46, indicating that L2 learners (M = 6.44, SD = 0.12) were no faster or slower than monolinguals (M = 6.48, SD = 0.19). There were no significant interactions (Level x Language status: F(1, 39) = 0.45, p = .5; Level x Word status: F(1, 39) = 0.2, p = .66; Language status x Word status: F(1, 39) = 0.2, p = .66; Level x Language status x Word status: F(1, 39) = 0.18, p = .67).

A 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) x 2 (word status: word, nonword) ANOVA was conducted on accuracy with level and language status as between-subject variables and word status as a within-subject variable. Words (M = .96, SD = .02) were no more or less accurate than nonwords (M = .95, SD = .02), F(1, 39) = 2.64, p = .11. There was not a main effect of level, F(1, 39) = 0.19, p = .66, indicating that advanced learners and matched monolinguals (M = .95, SD = .03) responded no more accurate than beginning learners and matched monolinguals (M = .95, SD = .02). There was also not a main effect of language status, F(1, 39) = 1.52, p = .23, indicating that L2 learners (M = .96, SD = .02) were no more accurate than monolinguals (M = .95, SD = .03). There was a significant interactions between level x language status, F(1, 39) = 6.29, p = .02, which reveals the same pattern of results as the overall d’ measure. Advanced L2 learners are more accurate than matched monolinguals, t(21) = 2.14, p = .04, but there are no other differences between groups. None of the other interactions were significant (Level x Word status: F(1, 39) = 0.23, p = .6; Language status x Word status: F(1, 39) = 0.39, p = .53; Level x Language status x Word status: F(1, 39) = 0.23, p = .63).
Statistical analyses were not performed on the ERP data of the matched subgroups, due to the very small number of participants per group (n = 11). However, the waveforms and scalp distributions will still be presented visually and discussed.

Figure 3-32: Accuracy rates for English words (red) and nonwords (blue) for advanced learners and monolinguals (left) and beginning learners and monolinguals (right). Advanced learners were more accurate than the matched monolinguals. There was not an effect of word status.
Figure 3-33: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -4 to 4) for learners and the matched monolingual groups showing differences.
between English words and nonwords.

Visual inspection reveals an N400 in all four groups, though slightly different magnitudes and scalp distributions. Interestingly, the advanced learners seem to detect the distinction between words and nonwords earlier than the other groups, as the N400 can be seen in the advanced learners’ scalp distributions appearing focally in the centro-parietal sites within the 300-400ms time window. The other three groups show a more left-lateralized effect, and the two beginning level groups show a more anterior effect. The monolinguals matched to the advanced learners show similar trends in terms of early registration of the distinction between words and nonwords, but the strength of the effect is not as pronounced as in the advanced learners. This coincides with the behavioral data that demonstrated that the advanced learners were more accurate than the matched monolinguals.

**Cognates vs. Noncognates**

A 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) x 2 (cognate status: cognate, noncognates) ANOVA was conducted on reaction times. There were no significant main effects (level: F(1, 39) = 2.04, p = .16; language status: F(1, 39) = 0.34, p = .56; cognate status: F(1, 39) = 0.94, p = .34), and there were no significant interactions (level x language status: F(1, 39) = 0.91, p = .35; level x cognate status: F(1, 39) = 1.05, p = .31; language status x cognate status: F(1, 39) = 0.13, p = .73; level x language status x cognate status: F(1, 39) < 0.001, p = .99).

A 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) x 2 (cognate status: cognate, noncognates) ANOVA was conducted on accuracy. There was a main effect of cognate status, F(1, 39) = 15.97, p < .001, revealing that participants were more accurate in responding to cognates (M = .97, SD = .28) than noncognates (M = .95, SD = .03). There was not
a main effect of level, F(1, 39) = 0.09, p = .77, nor was there a main effect of language status, F(1, 39) = 0.2, p = .65. There were no significant interactions (level x language status: F(1, 39) = 2.38, p = .13; level x cognate status: F(1, 39) = 0.48, p = .49; language status x cognate status: F(1, 39) = 0.51, p = .48; level x language status x cognate status: F(1, 39) = 0.12, p = .74).

Looking at the ERPs for the different groups suggests that, despite the lack of behavioral differences between cognates and noncognates, the L2 learners are showing some sensitivity to the presence of English-Spanish cognates.
Figure 3-34: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -2 to 2) for learners and matched monolinguals showing differences between English cognates and noncognates.

The advanced learners’ waveform displays a persistent reduced negativity for cognates compared to noncognates, and the scalp distribution suggests that the effect is, in fact, an N400. The matched monolinguals are showing a small (~1 microvolt) effect of something in the anterior electrode sites, but the distribution does not lead to the interpretation that it is an N400. The beginning learners, however, are also showing a small (~1 microvolt) effect, but in the electrode sites where an N400 is typically found. Looking at their waveforms reveals that the cognate/noncognate distinction may actually be processed slightly later than the standard 300-
500ms time window. An additional scalp plot from 550-650ms reveals what appears to be a delayed N400 effect in the predicted direction.

![Scalp Distribution](image)

Figure 3-35: Scalp distribution for the beginning learners from 550-650 ms, revealing a small (~1 mv) effect for cognates in centro-parietal regions consistent with an N400.

Together, these data suggest that L2 learners are, in fact, sensitive to language overlap while processing in the L1, despite the fact that the sensitivity does not manifest yet in behavior. Whereas all four groups, including monolinguals, were more accurate for cognate words than noncognates words (which may contribute to the anterior effect found in the monolinguals matched to the advanced learners), only the L2 learners display the patterns of neural effects in line with an N400 effect. The beginning learners appear to access the L2 information later than the advanced learners, leading to a slightly delayed N400 effect in the L1 for cognate words.

**Homographs vs. Controls**

A 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) x 2 (homograph status: homograph, control) ANOVA was conducted on reaction times. There were no significant main effects (level: F(1, 39) = 2.55, p = .12; language status: F(1, 39) = 0.4, p =
.53; homograph status: F(1, 39) = 0.14, p = .71), and there were no significant interactions (level x language status: F(1, 39) = 0.32, p = .57; level x homograph status: F(1, 39) = 0.31, p = .58; language status x homograph status: F(1, 39) = 0.13, p = .72; level x language status x homograph status: F(1, 39) = 0.68, p = .41).

A 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) x 2 (homograph status: homograph, control) ANOVA was conducted on accuracy. There was a main effect of homograph status, F(1, 39) = 8.78, p = .01, revealing that participants were more accurate in responding to homograph controls (M = .97, SD = .03) than homographs (M = .94, SD = .05). There was not a main effect of level, F(1, 39) = 0.72, p = .4, nor was there a main effect of language status, F(1, 39) = 0.6, p = .44. There were no significant interactions (level x language status: F(1, 39) = 3.35, p = .08; level x cognate status: F(1, 39) = 1.95, p = .17; language status x cognate status: F(1, 39) = 1.42, p = .24; level x language status x cognate status: F(1, 39) = 2.55, p = .12).

The ERP data for the homographs and control words suggests that the differences in accuracy between homographs and controls may be due to a factor other than English-Spanish overlap, since the matched monolinguals show similar patterns to the L2 learners and the distribution of the effect is unlike the distribution predicted by an N400.
Figure 3-36: Waveforms from electrode Cz, and scalp distributions from 300-400 ms and 400-500 ms (scale: -2 to 2) for learners and matched monolinguals showing differences between English homographs and controls.
The advanced learners and matched monolinguals appear to show a frontal central negativity for homographs compared to controls. This effect is especially pronounced in the matched monolinguals, whose effect is upwards of two microvolts. The effect appears in the earlier time window for the advanced monolinguals and only in the later time window for the advanced learners.

The beginning learners and matched monolinguals, on the other hand, appear to show a completely different effect. Both groups demonstrate less negativity for homographs compared to control words, unlike the advanced groups, and the effect appears in posterior electrodes and differentially lateralized in different time windows. Also unlike the advanced groups, the beginning learners show the effect emerging in the earlier time window, and the matched monolinguals only in the later time window.

The Cz electrode was displayed for the purpose of comparison across groups and across other conditions, since it has been used throughout these results, but to better capture the effects seen in the scalp distributions, ERPs from electrode Fz for the advanced groups and ERPs from the electrode P4 for the beginning groups are displayed below.

Figure 3-37: Waveforms for advanced learners and matched monolinguals from an anterior electrode (Fz) showing increased negativity for homographs compared to control words in both groups.
The magnitude of the effect is apparent in these waveforms at more frontal sites, as well as the fact that the differences are most likely not an N400. The difference between homographs and controls appears very early in the matched monolinguals and persists throughout the epoch, and though it appears later for the advanced learners, once it differentiates, it also persists throughout the epoch.

Figure 3-38: Waveforms for beginning learners and matched monolinguals from a right-lateralized posterior electrode (P4), revealing a similar magnitude of reduced negativity for homographs and control words.

The magnitude of the effect in the beginning groups is much smaller, but the homographs are consistently less negative than the homograph control words. It is interesting to note that the difference in the homograph effect appears between the level (beginning vs. advanced) rather than language status (learner, monolingual). The four participant groups here do not differ in any measured demographic factors (age, gender, self-rated L1 proficiency) or cognitive factors (O-Span, Flanker Effect, AX-CPT), so it is unclear what the difference between beginning and advanced participants.

The behavioral data, which found that participants were less accurate on homograph trials than controls words, is better characterized by the advanced participants’ ERP data (both learners and monolinguals), which showed greater negativity for homographs compared to controls,
consistent with interference. The magnitude of the effect in the beginning group was much smaller, but the trend toward facilitation is puzzling.

**Verbal Fluency**

A 2 (level: advanced, beginning) x 2 (language status: learner, monolingual) ANOVA was conducted on the number of exemplars produced in the English verbal fluency task. The advanced group (M = 18.53, SD = 2.98) and the beginning group (M = 20.26, SD = 3.26) did not differ in the amount of English exemplars they produced, F(1, 33) = 2.6, p = .12. The L2 learners (M = 19.09, SD = 3.68) and the monolinguals (M = 20, SD = 2.75) also did not differ in the number of English exemplars, F(1, 33) = 0.65, p = .43. There was not an interaction between level and language status, F(1, 33) = 1.92, p = .18.

For the Spanish verbal fluency task, a t-test was conducted to compare the advanced and beginning L2 learners. Advanced learners (M = 8.26, SD = 1.73) produced more responses in Spanish than the beginning learners (M = 4.84, SD = 3.18), t(16) = 3.11, p < .01.

**Continuous Statistics**

For the purposes of evaluating the relationship between L2 proficiency, sensitivity to the L2 during L1 processing, and L1 costs, correlational analyses were conducted on the L2 learners only. Monolingual data were excluded from all correlations presented in this section.

An L2 proficiency measure was created using the verbal fluency performance in each language: average number of L2 exemplars produced / average number of L1 exemplars produced. This type of ratio proficiency measure has been used in past research (Pivneva, Mercier, & Titone, 2014) and was selected for the analyses here for several reasons. First, this
proficiency measure consists of data from production tasks, which appear to be particularly sensitive to L2 proficiency, and semantic fluency has been associated with vocabulary size in past research (Luo, Luk, & Bialystok, 2010). Second, this proficiency measure controls for L1 fluency by including it in the denominator. One of the persistent findings in the present study, as well as past research (Sparks, Ganschow, & Patton, 1995; Sparks et al., 1998), was that participants who demonstrated higher English (L1) proficiency also had higher Spanish (L2) proficiency (see Figure 3-39). This proficiency measure accounts for general language fluency but also provides a sense of the balance between both languages, where higher values indicate that the participant’s Spanish proficiency is approaching his/her English proficiency.

Figure 3-39: Relationship between L2 proficiency (L2/L1 verbal fluency measure) and English performance (d’) reveals a positive correlation ($r = .38$, $p = .02$). Higher L2 proficiency is related to higher L1 proficiency.
In line with the relationship between L1 and L2 proficiency is the finding that English d’ and Spanish d’ are correlated ($r = .67, p < .001$) and English RT and Spanish RT are correlated ($r = .7, p < .01$). In both languages, participants who are faster overall are also faster on nonwords, making the difference between words and nonwords (hereafter the word effect) smaller.

The two effects used to index L2 sensitivity during L1 processing are the cognate effect and homograph effect. The cognate effect is cognates minus noncognates, and the homograph effect is homographs minus control words. For reaction times, facilitation yields negative values (since faster reaction times minus slower reaction times are negative) and interference yields positive values.

Looking at the relationship between the cognate and homograph effect in English reveals that participants who show facilitation for cognates also show interference for homographs ($r = -.28, p = .06$), as seen in Figure 3-40. However, this pattern does not remain true on the ERP record. A correlation between the magnitude of the N400 for cognates and homographs reveals that when cognates are less negative than noncognates on the ERP record, homographs are also less negative than controls ($r = .3, p = .05$), as in Figure 3-41.
Figure 3-40: Relationship between English cognate effect and English homograph effect ($r = -.28$, $p = .06$). Behaviorally, these effects are negatively correlated.

Figure 3-41: Relationship between English cognate effect and English homograph effect on the ERP record shows a positive relationship ($r = .3$, $p = .05$).
Given the strong correlation between L1 and L2 proficiency, where higher scores in English correlated with higher scores in Spanish, there was little evidence to support the hypothesis that L2 learners who endure costs to the L1 achieve higher L2 proficiency. However, there was one relationship that may be related to this prediction. The relationship between the cognate effect in English, a measure of L2 sensitivity, and English d’ scores revealed that participants who displayed greater interference for English cognates (i.e., more positive values) also had lower English d’ scores ($r = -.34, p = .02$). However, the reverse is also true, such that participants who were facilitated by English cognates had higher English d’ scores. While cognates typically facilitate processing, they also sometimes produce interference, which is most commonly found in the L1 (Kroll et al., 2002). In this case, some learners were facilitated in the L1 while others experienced interference. What characteristics about a learner or a context modulate these effects is yet unknown.

Figure 3-42: Relationship between L2 sensitivity (English cognate effect) and English
The other prediction made by the hypothesis tested here is that learners who demonstrate sensitivity to the L2 during L1 processing should have higher L2 proficiency. Evidence for this prediction comes from looking at the relationship between the English cognate effect and the Spanish word effect. Participants who are better able to distinguish Spanish words and nonwords, as indexed behaviorally by the magnitude of the word effect and on the ERP record via the magnitude of the N400, should be more proficient in Spanish. Participants who had a larger N400 showed greater facilitation for English cognates \((r = -.51, p = .02)\) (Figure 3-43), and participants who had a larger (i.e. more negative) word effect behaviorally also showed marginally greater facilitation for English cognates \((r = .28, p = .06)\) (Figure 3-44).

Figure 3-43: Relationship between Spanish proficiency, as operationalized by the word vs. nonword distinction on the ERP record, and L2 sensitivity, as operationalized by the English cognate effect. Greater English cognate facilitation is associated with a larger N400 for Spanish words vs. nonwords.
Against expectations, none of the individual difference measures were correlated with each other, or with language processing variables. The flanker effect was not correlated with the BSI index ($r = .17, p = .27$) or with O-Span scores ($r = 0, p = .98$), nor were the BSI and O-Span scores correlated ($r = .04, p = .78$). The flanker effect did not correlate with any of the English or Spanish language variables (range of $r$: -.12 to .16) and neither did the BSI (range of $r$: -.08 to .23). The O-Span scores were positively correlated to English $d'$ ($r = .38, p = .01$) and Spanish $d'$ ($r = .28, p = .07$). However, the strong relationship between English and Spanish $d'$ scores independent of O-Span scores led to a mediation analysis. The mediation analysis tested the idea that working memory positively influences L1 fluency, which in turn promotes L2 fluency. Using Baron and Kenny’s (1986) steps for a mediation analysis, it was found that working memory (O-
Span scores) predicts L2 proficiency (Spanish d’), $\beta = 0.005$, $p = .01$. Working memory also predicts L1 proficiency (English d’), $\beta = 0.01$, $p = .002$. When working memory and L1 proficiency are jointly used as predictors of L2 proficiency, the relationship between working memory and L2 proficiency disappears, $\beta = .001$, $p = .51$, whereas the relationship between L1 and L2 proficiency is strongly maintained, $\beta = .41$, $p < .001$, thereby satisfying the requisites for a mediation relationship (Baron & Kenny, 1986). This suggests that working memory plays an important role in language processing in general, but is most crucial in the development of L1 fluency, promoting overall language aptitude and L2 learning.
Chapter 4  
Discussion

The purpose of this study was to test a new hypothesis about second language learning that predicts that L2 learners differ in their tolerance to changes in the L1, which may be a necessary step during L2 learning, and that those who are better able to tolerate L1 changes should achieve higher L2 proficiency. To test this hypothesis, L2 learners at various levels of Spanish proficiency performed a lexical comprehension and production task in English (L1) and Spanish (L2) and were compared to matched monolingual English participants on tasks in English. Critically, ERPs were used in the lexical comprehension task, which revealed emerging effects within the groups of L2 learners that were statistically absent behaviorally, which suggested that their L2 knowledge was active during L1 processing. The effect within the advanced learners was striking, while the effect within beginning learners was smaller but nonetheless surprising.

The predictions made by the hypothesis regarding L2 sensitivity during L1 processing garnered mixed support from the data reported here. Behaviorally, there were no differences in reaction times between cognates and noncognates, or homographs and controls. Overall, participants responded more accurately to cognates compared to noncognates, and less accurately to homographs compared to controls, but neither of these differences interacted with group; monolinguals were also more accurate for cognates and less accurate for homographs, suggesting item variability.

However, the ERPs revealed sensitivity that did not manifest in behavior. Overall, cognates were less negative in the 300-500ms time window compared to noncognates. Although
this effect did not interact with group, the visual appearance of the waveforms and scalp
distributions suggests that the advanced and beginning learners showed an emerging N400 effect
(with cognates as less negative than noncognates), whereas the monolinguals showed little
reduced negativity for cognates in very anterior sites unlike the distribution of the N400. These
ERP data support the prediction that L2 learners show sensitivity to their L2 knowledge while
using the L1. Though the effect was not significant for the beginning learners, the emerging effect
was visually present in the waveform and particularly in the scalp distribution. The finding of L2
sensitivity in the L1 of the advanced learners is particularly important in characterizing the
trajectory of L1 changes as L2 proficiency increases. To date, not many studies have found
evidence of L2 sensitivity during L1 processing on the ERP record. Although the advanced
learners were characterized as such, they were still L2 learners rather than proficient bilinguals,
and their L2 self-ratings were only 6.9 on a 1-10 scale.

One of the past studies finding a cognate effect in the L1 was conducted by Midgley,
Holcomb, and Grainger (2011). They asked highly proficient L2 learners to perform a semantic
decision task, which was to decide if each word was an animal or not. Among the items were
cognates and noncognates. Participants performed the task in the L1 and the L2. Like the present
results, they found a reduced N400 for cognate words compared to noncognate words in both
languages. The advanced learners in the present study revealed a cognate effect in both English
and Spanish (discussed below). However, past research has not investigated when the cognate
effect emerges as a function of L2 proficiency. The present results suggest that even beginning L2
learners, with as few as two semesters of language classes, begin to show neural sensitivity to the
presence of language ambiguous words.

The effect of homographs was a little less clear. The advanced learners demonstrated
increased negativity for homographs compared to controls, but in central anterior sites rather than
centro-parietal sites like an N400. The beginning learners, on the other hand, did show a small
central reduced negativity for the homographs, slightly left-lateralized, which may or may not be an N400. However, given the behavioral finding that homograph words were less accurate than the control words, the expectation would be to find increased negativity for homographs relative to controls, rather than reduced negativity as found in the beginning learners. Finally, the monolinguals demonstrated persistent negativity for homographs compared to controls in the waveforms, but again, the negativity appears in frontal sites close to the eye electrodes.

I hesitate to interpret the effect of homographs within the context of the predictions made by the hypothesis since the monolinguals appear to be somewhat sensitive to these homographs too. On the one hand, within the group of L2 learners, the magnitude of the behavioral cognate effect and homograph effect in English were negatively correlated, meaning that L2 learners who were facilitated by cognates showed interference for homographs and vice versa. The magnitude of the ERP cognate and homograph effects were positively correlated, showing the opposite pattern, that participants who had reduced negativity for cognates also had reduced negativity for homographs. This suggests some relationship between cognate and homograph processing within the L1. On the other hand, the fact that the monolinguals showed an effect is puzzling. I do not challenge the monolingual status of these participants, since they were not sensitive to the cognates, which showed a more clearly predicted pattern of effects and which are generally more sensitive to L2 experience. However, it may be that a lexical characteristic was overlooked within the English materials, creating another difference between the homographs and controls that produce the divergent effects found across these groups that is not attributed to language knowledge. Future analyses will include an item analysis of the English homograph data to look only at the homograph and control words for which monolinguals showed no effect. Looking at that subset of items within the beginning and advanced learners will illuminate whether they show an effect of homograph status in the L1.
The continuous statistics also provide some insight into the effects of the L2 on the L1. Looking once again to the magnitude of the behavioral cognate effect in English, it appears that participants who were more facilitated by the cognates in English were also the participants who demonstrated the greatest sensitivity to the distinction between Spanish words and nonwords. This would seem to support the prediction that greater L2 sensitivity during L1 processing is related to higher L2 proficiency. However, in this instance, the reverse is also true. Participants who showed interference in response to English cognates also showed a comparatively smaller N400 magnitude for Spanish words and nonwords. Experiencing facilitation or interference in response to English cognates should be evidence for L2 sensitivity; the reason why cognate facilitation would result in higher L2 proficiency but not cognate interference is interesting and remains unanswered. However, past research has found cognate interference in the L1 and cognate facilitation in the L2 among both lower proficiency and higher proficiency L2 learners for word naming (Kroll et al., 2002). This further complicates the reported relationship between cognate facilitation and L2 proficiency, but also confirms that L1 cognate interference is not an entirely unreported effect.

The other prediction made by the hypothesis tested here is that L2 learners who demonstrate costs in the L1 will achieve higher L2 proficiency. This prediction found little to no support in the data. The beginning learners did not differ from the monolinguals as a group in any of the tasks or measures, including when the groups were matched on demographic and cognitive factors. When the advanced learners differed from the matched monolinguals in English d’, they proved to be more accurate rather than less accurate. Finally, one of the most persistent findings when conducting the continuous statistics was that higher L1 proficiency is related to higher L2 proficiency. One example of this was the strong positive correlation between the L2 proficiency measure and English d’, indicating that higher Spanish proficiency is related to higher accuracy in English. This finding has been established and replicated in past research, which has also found
that grades in English classes and scores on a language aptitude test predict grades in second
language courses (Sparks et al., 1995), and that participants higher in English proficiency differed
from participants lower in English proficiency on several measures of L2 proficiency after two
years of foreign language classes (Sparks et al., 1998). As mentioned in the introduction, the L1
costs were predicted to be greater in the beginning learners than the advanced learners due to the
trajectory of these costs found in past literature, which was not supported.

The only potential support for this prediction comes from the correlation between the
English behavioral cognate effect and English d’. Participants who demonstrated interference for
English cognates also had lower English d’ scores. If cognate interference in English indexes L2
sensitivity, then this would suggest that greater L2 sensitivity is related to worse English
accuracy. However, this argument is tenuous, at best, for various reasons. As mentioned,
participants who demonstrated greater facilitation for English cognates also demonstrated the
greatest sensitivity to the Spanish word/nonword distinction, and vice versa. That is, participants
who demonstrated interference for English cognates demonstrated the least sensitivity to Spanish
words. Combining this information with the first correlation suggests that participants who
demonstrate interference for English cognates are less sensitive to the word/nonword distinction
in Spanish and also less accurate in English. Once again, we encounter the persistent finding that
L1 and L2 proficiency are strongly related. The group-level analyses did not reveal any statistical
support for the presence of L1 costs, but the boxplots of performance did reveal that the is vast
variability within the beginning learners in terms of performance, many of whom were
performing worse than the monolinguals. Still, this did not manifest in any of the continuous
statistics.

The fact that this prediction was not supported within the data presented here is not
terribly surprising. The phrasing of the prediction is that L1 costs (in the present) will lead to
higher L2 proficiency (in the future). The concept of desirable difficulty is that you must pay the
price now to reap the benefits later on. Therefore, the design of this study was not ideal to detect
evidence for this prediction since the design was cross-sectional in nature and the prediction
requires the tracking of L1 and L2 performance over time. One future avenue of research is to
conduct a mini-longitudinal study on beginning L2 learners over the course of one to two
semesters of L2 learning. This would track L1 performance over time, to detect if the beginning
stages of L2 learning impose costs on the L1, which can then rebound in fluency by a later time
point. This design would also permit the L2 proficiency to be measured at two time points, to
determine the rate and level of L2 attainment.

However, another interesting explanation for the lack of L1 costs draws upon another
assumption within the hypothesis. If we conceptualize L1 costs as L1 inhibition, then this
promotes the separation of the L1 and L2 to allow the L2 to grow and should appear primarily in
the beginning stages of learning. With time and proficiency, the L2 should come to affect the L1,
but L1 costs and L2 sensitivity should not be concurrent, and L1 costs should be relatively
unrelated to L2 proficiency. The alternative conceptualization is that L1 costs are the result of the
L1 facing new and uncharted competition from a new language (L2). This is entirely possible,
due to the fact that even beginning learners were showing emerging effects of cognates within the
L1. L1 costs would then be concurrent and related to L2 sensitivity. The data reported here do not
necessarily support either of these explanations, due to the lack of evidence for L1 costs, but also
do not preclude the possibility of either.

Perhaps some of the most interesting data in the present study came from the Spanish
lexical decision task. The data from Spanish words and nonwords showed the expected pattern.
Words elicited faster reaction times and higher accuracy, and the advanced learners were faster
and more accurate than the beginning learners. The ERP data concur, revealing that both groups
had reduced negativity for Spanish words compared to nonwords. The scalp distributions are not
picturesque N400s, as both groups showed the effect in more anterior regions, but by the later time window the advanced learners showed a distribution more like the classic N400.

The surprising data come from the Spanish cognates and homographs. The behavioral data from the Spanish cognates revealed that cognates elicited faster and more accurate responses, and that advanced learners were faster and more accurate. However, the ERP data showed that the beginning learners had no reduced negativity for the cognates compared to noncognates. The mean amplitude graphs showed almost equal mean amplitudes for Spanish cognates and noncognates in the beginning learners, their waveform demonstrated overlapping lines, and the scalp distribution revealed no effect. The advanced learners demonstrated the predicted effects, with reduced negativity for cognates, and an N400 in the waveform and scalp distribution. The beginning learners benefitted equally from the presence of cognates in the behavioral data given the lack of a group x cognate status interaction, so why did the ERP data not showing this sensitivity?

One explanation is that the beginning learners had such little proficiency in Spanish that they saw the cognates and interpreted them as English. More than one participant clarified during the Spanish lexical decision task that they should only push the button for words if the item was a word in Spanish, not English. One might expect that these participants saw the cognates and assumed they were English words. However, this cannot be true given that the beginning learners were faster and more accurate for cognates. If they interpreted the words as English words, then the ERPs also should have detected greater sensitivity or the processing of conflict between languages rather than no sensitivity.

Another possibility returns to the observation that the beginning learners had very high variability in accuracy and reaction times in Spanish. Though expected, it also suggests a high level of guessing behaviors, which would effectively wash out the effects of cognates. This makes comparing beginning and advanced learners complicated, since the advanced learners are
correctly answering more trials correctly, with a broader range of lexical characteristics (lower frequency words, longer words), whereas the correct responses for beginning learners include correct guesses (fast reaction times, any range of lexical characteristics) as well as correct responses for known words (slower reaction times, most likely short and high frequency words). Therefore, one future analysis will evaluate the subset of Spanish items that beginning learners consistently answered correctly and comparing those items across beginning and advanced learners. This may reveal a cognate effect in Spanish that was eliminated due to the higher variability in their responses.

These results also conflict with past research on the cognate effect as a function of proficiency. Tonzar, Lotto, and Job (2009) taught an L2 and L3 to fourth- and eight-grade students via word-translation or picture-translation. They included some words in each language that were cognates and others that were noncognates. Regarding the cognate results, they found that across groups, the magnitude of the cognate effect was largest at the first testing session and decayed as proficiency in the L2 and L3 increased over time. The results of the present study demonstrate the opposite pattern, though on a different time scale. The beginning learners showed the smallest neural sensitivity to cognates whereas the advanced learners showed a clear cognate effect on the ERP record. Behaviorally, the results from Tonzar et al. (2009) would predict an interaction between group and cognate status to reveal that the effect was larger in beginning learners than advanced learners, but this was also not found.

The Spanish homographs had a similarly perplexing pattern of results. The potential item-level correlations present within the English homograph data do not apply here, since the materials were not counterbalanced across languages. Both groups were more accurate for homograph words than controls, though this difference was more defined in the advanced learners due to the large variance in the beginning learners’ accuracy rates. The ERP data also reveal an interaction between group, homograph status, and anteriority, in which the advanced learners
have increased negativity for homographs across all sites, but particularly in the posterior regions. The beginning learners, on the other hand, had very small but opposing effects in anterior and posterior regions, demonstrating reduced negativity (i.e. facilitation) for homographs in posterior sites (especially the occipital electrodes) and increased negativity in anterior sites (i.e. interference) like the advanced learners. Based on the behavioral data that showed higher homograph accuracy, we would expect to see facilitation in the neural response. The homographs, like the cognates, reveal a dissociation between behavioral and neural processing.

Past research has also found this dissociation between neural and behavioral data. McLaughlin et al. (2004) asked beginning learners of French to perform a lexical decision task in French in which each real word was preceded by a related or unrelated word. Although the learners were no better than chance at discriminating words and nonwords behaviorally, their ERPs showed an N400 for words compared to nonwords after 14 hours of instruction, and a further distinction between related and unrelated word pairs after 63 hours of instruction. This study suggests that the direction of motion is from neural sensitivity to behavioral sensitivity. This may be true for the English lexical decision task, in which neural sensitivity to the cognates was found within both groups of learners, though stronger in the advanced learners, but had not yet appeared in the behavioral data. However, this is not the same for the Spanish lexical decision task, in which the cognates and homographs produced behavioral effects that were not present in the ERPs.

The role of transfer was described briefly in the introduction, at which point the prediction was made that L2 learners who are effectively able to reduce the influence of the L1 during L2 processing (i.e. who transfer less L1 knowledge) should be more proficient in the L2. This prediction was falsified here with the results from the Spanish cognates in beginning and advanced learners, which suggest that higher proficiency reveals greater L1 influence. Interestingly, the evidence also suggests that the L2 knowledge transferred to the L1 in the
beginning learners before the L1 knowledge transferred to the L2. This is very unexpected, especially given that researchers hadn’t considered the idea that L2 knowledge could transfer back to the L1 until recently. The majority of the research on transfer has looked at the level of the grammar, thus the fact that the present study investigated the lexical level may account for the different findings.

Finally, it is worth noting the lack of any effects found with the individual difference measures. Neither the Flanker effect nor the AX-CPT scores differed across groups or were related to any continuous measures of proficiency or L1 performance. This may suggest that inhibitory control does not play a role in the development of L2 proficiency, the regulation of the L1, or the rate and magnitude of transfer. However, an equally likely explanation is that our current measures of inhibitory control are not sensitive enough to distinguish between these groups of young participants who are at their peak cognitive functioning. The prediction involving the role of inhibitory control was that inhibitory control would modulate L1 costs, such that L2 learners with higher inhibitory control would be better able to inhibit the L1. Given that there was little to no evidence for the prediction of L1 costs, it may also be that the role of inhibition was masked.

Several approaches will be taken in future analyses to better characterize the involvement of inhibitory control. Each task used (flanker, AX-CPT) produces a single value for the general effect (flanker effect, BSI), but also contain several sub-components that may relate to different aspects of inhibitory control. For example, Morales et al. (2013) found that error rates in the AY condition of the AX-CPT task were correlated with the stop-signal reaction times, and that the difference between reaction times in AY and BY trials were related to overall response times, and that these effects differed by language status (monolinguals vs. bilinguals). Therefore, these subcomponents will be analyzed in relationship to other inhibitory measures as well as language processing measures. One particular relationship of interest is between AY trials in the AX-CPT
task, which indexes reactive inhibitory controls, and false alarm rates in the lexical decision task. False alarm rates are more likely related to reactive compared to proactive inhibitory control, so this may reveal how false alarm rates are modulated by a participant’s reactive inhibitory control.

Some of the limitations of the present study have already been discussed, such as the cross-sectional rather than longitudinal design of the study. Another limitation is the number of participants. The use of ERPs, and especially the matching of a subset of learners and monolinguals, would have benefitted from a larger number of participants per group. Many of the ERP effects were visually present but statistically absent, indicating a lack of power. Finally, it would have been beneficial to include a single task with the purpose of evaluating L2 proficiency, preferably one that measures L2 grammatical proficiency. Vocabulary size is a good proxy for L2 proficiency, but the present study is unable to speak to the relationship between lexical and syntactic proficiency. Furthermore, the proficiency measure used here, the ratio of L1 to L2 verbal fluency exemplars, precludes the option of evaluating the relationship between L2 proficiency and L2 production levels since the predictor (L2/L1 verbal fluency) would contain the dependent measure (L2 verbal fluency, L1 verbal fluency).

Many more analyses can be conducted on the data collected in the present study. Of interest is a multilevel model of the lexical decision data. For the English lexical decision task, both monolinguals and L2 learners will be included to evaluate the joint contribution of cognitive and inhibitory control measures, and the role of L1 proficiency measured via the verbal fluency task. A separate model will be created for the L2 learners only that will include English Spanish lexical decision data, to investigate how L2 sensitivity may differentially appear across languages or levels of proficiency.

Despite the limitations, the present study does provide valuable information regarding the nature of L1 changes at different levels of L2 proficiency. As in past research, the advanced learners demonstrated sensitivity to the cognates in L1 processing, but the present study
additionally found emerging evidence in much less proficient L2 learners. The research finding effects of the L2 on the L1 in L2 learners is limited, thus the present study contributes to this growing body of research and further provides an idea of the trajectory of this sensitivity. The role of L1 transfer is qualified by the findings in the present study, given that the beginning learners did not show neural sensitivity to cognates in the L2 but did in the L1, though future analyses may clarify this non-finding. The strong correlation between L1 and L2 fluency provides an important factor for teachers and students of second language classes to keep in mind for evaluation and improvement. The present study provides a glimpse into what happens to the native language during second language learning.

Future avenues of research should consider a longitudinal study to evaluate changes to the L1 and the L2 over time. Additionally, the present study only looked at changes to the lexicon, while future studies should investigate changes to the grammar. Past research has looked at changes to the phonology, but within immersed learners (Chang, 2013), thus an investigation of the changes to phonology within classroom learners would also be fruitful.
References


Appendix A

Language History Questionnaire

This questionnaire is designed to give us a better understanding of your experience with other languages. We ask that you be as accurate as thorough as possible when answering the following questions.

Part I

1. Gender: ___________________
2. Age: ______ years
3. Native Country
   - United States
   - Other [Please specify: ___________________
4. Native Language (Please check all that apply)
   - English
   - Other [Please specify: ___________________
5. Language(s) spoken at home. (Please check all that apply)
   - English
   - Spanish
   - German
   - Chinese
   - Other [Please specify: ___________________

Part II

The next section of the questionnaire deals with your second language learning experience.

6. Have you studied any second language(s)?
   - No → If NO, please go to Part IV (on the final page of this questionnaire)
   - Yes → If yes, which language(s)?
____________________________________________________________

7. Where and when did you study a second language before college? Please check all that apply and indicate length of study. If you have studied more than one language at any given time, please provide information about each language you have learned.
   - Home – Language(s): ____________________________ Since what age? ________
Elementary School – Language(s): _______________________________ For how many years? __________

Middle School – Language(s): _______________________________ For how many years? __________

High School – Language(s): _______________________________ For how many years? __________

8. Have you studied any second language(s) in college?
   - No → If NO, please go to Question # 10
   - Yes → If yes, which language(s)?

   For how long?
   - Less than a one semester
   - 1-2 semesters
   - 3-4 semesters
   - 5-6 semesters
   - 7-8 semesters
   - 8+ semesters

9. Are you: (Please check all that apply and indicate which language each applies to each if you have studied more than one second language at college.)
   - A Spanish, German, etc. student
   - Taking a second language for a requirement but interested in being a major or minor.
   - Taking a second language for a requirement; NOT interested in being a major or minor.
   - A second language minor
   - A second language major
   - A second language graduate student
   - Other [Please explain: ________________________________]

10. Have you studied and/or lived abroad?
    - Yes
    - No

   If Yes, where and when did you study, for how long, and what language(s) did you speak?

<table>
<thead>
<tr>
<th>Country</th>
<th>Approx. dates</th>
<th>Length of Stay</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11. What do you consider to be your primary second language? (You may check more than one if you feel that you have multiple “primary” second languages.)
    - English
12. What language do you currently think is your *dominant* language (i.e., the language you are most comfortable using on a daily basis)? (Please check one)

- Spanish
- German
- Chinese
- Other [Please specify: ________________________________]
Part III

The next section asks you to rate your skills in your Spanish.

13. Your second language reading proficiency. (1=not literate and 10=very literate)

   1  2  3  4  5  6  7  8  9  10

14. Your second language spelling proficiency. (1=not good and 10=very good)

   1  2  3  4  5  6  7  8  9  10

15. Your second language writing proficiency. (1=not literate and 10=very literate)

   1  2  3  4  5  6  7  8  9  10

16. Your second language speaking ability. (1=not fluent and 10=very fluent)

   1  2  3  4  5  6  7  8  9  10

17. Your second language speech comprehension ability. (1=unable to understand conversation and 10=perfectly able to understand)

   1  2  3  4  5  6  7  8  9  10

18. In my second language classes I get:

   - Mostly A's
   - Mostly A's and B's
   - Mostly B's
   - Mostly B's and C's
   - Mostly C's

19. Please explain any additional language experience that you have had. Please indicate how many years you studied the language, your level of proficiency, etc.
Part IV
The next section of the questionnaire deals with your English language skills. Please rate yourself on each measure by circling the appropriate number.

These ratings are for English.

20. Your English reading proficiency. (1=not literate and 10 = very literate)
   1 2 3 4 5 6 7 8 9 10

21. Your English spelling proficiency. (1=not good and 10=very good)
   1 2 3 4 5 6 7 8 9 10

22. Your English writing proficiency. (1=not literate and 10=very literate)
   1 2 3 4 5 6 7 8 9 10

23. Your English speaking ability. (1=not fluent and 10=very fluent)
   1 2 3 4 5 6 7 8 9 10

24. Your English speech comprehension ability. (1=unable to understand conversation and 10=perfectly able to understand)
   1 2 3 4 5 6 7 8 9 10

Thank you for your participation!
Appendix B

English & Spanish Lexical Decision Task Materials
<table>
<thead>
<tr>
<th>Cognates</th>
<th>Noncognates</th>
<th>Spanish Cognates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target</strong></td>
<td><strong>Length</strong></td>
<td><strong>Frequency</strong></td>
</tr>
<tr>
<td>brutal</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>canal</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>chocolate</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>combination</td>
<td>11</td>
<td>57</td>
</tr>
<tr>
<td>crude</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>declare</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>delicious</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>disgust</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>doctor</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>exaggerate</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>general</td>
<td>7</td>
<td>497</td>
</tr>
<tr>
<td>generous</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>global</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>goal</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>gracious</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>habitual</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>hospital</td>
<td>8</td>
<td>110</td>
</tr>
<tr>
<td>imaginary</td>
<td>9</td>
<td>17</td>
</tr>
<tr>
<td>insist</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>instruction</td>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td>literature</td>
<td>10</td>
<td>133</td>
</tr>
<tr>
<td>manual</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>material</td>
<td>8</td>
<td>174</td>
</tr>
<tr>
<td>musical</td>
<td>7</td>
<td>85</td>
</tr>
<tr>
<td>ordinary</td>
<td>8</td>
<td>72</td>
</tr>
<tr>
<td>organize</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>original</td>
<td>8</td>
<td>103</td>
</tr>
<tr>
<td>participate</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>particular</td>
<td>10</td>
<td>179</td>
</tr>
<tr>
<td>piano</td>
<td>5</td>
<td>38</td>
</tr>
<tr>
<td>probable</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>professor</td>
<td>9</td>
<td>57</td>
</tr>
<tr>
<td>recipient</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>similar</td>
<td>7</td>
<td>157</td>
</tr>
<tr>
<td>superior</td>
<td>8</td>
<td>46</td>
</tr>
<tr>
<td>usual</td>
<td>5</td>
<td>96</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>7.69</strong></td>
<td><strong>61.33</strong></td>
</tr>
<tr>
<td>Homographs</td>
<td>Homograph Controls</td>
<td>Spanish Homographs</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td><strong>Target</strong></td>
<td><strong>Length</strong></td>
<td><strong>Frequency</strong></td>
</tr>
<tr>
<td>actual</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>agenda</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>arena</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>assistance</td>
<td>10</td>
<td>87</td>
</tr>
<tr>
<td>bark</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>bigot</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>cargo</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>carpet</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>cart</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>casual</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>code</td>
<td>4</td>
<td>40</td>
</tr>
<tr>
<td>contest</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>cure</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>dose</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>eleven</td>
<td>6</td>
<td>40</td>
</tr>
<tr>
<td>eventual</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>exit</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>fabric</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>former</td>
<td>6</td>
<td>131</td>
</tr>
<tr>
<td>grade</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>hill</td>
<td>4</td>
<td>72</td>
</tr>
<tr>
<td>journal</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>lecture</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>meter</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>mole</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>notice</td>
<td>6</td>
<td>59</td>
</tr>
<tr>
<td>pan</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>parade</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>pass</td>
<td>4</td>
<td>89</td>
</tr>
<tr>
<td>pie</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>pretender</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>realize</td>
<td>7</td>
<td>69</td>
</tr>
<tr>
<td>rope</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>salt</td>
<td>4</td>
<td>46</td>
</tr>
<tr>
<td>tea</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>torment</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>tuna</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>vine</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>5.32</strong></td>
<td><strong>29.34</strong></td>
</tr>
<tr>
<td>Target</td>
<td>Length</td>
<td>Frequency</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>-----------</td>
</tr>
<tr>
<td>abuso</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>actor</td>
<td>5</td>
<td>119</td>
</tr>
<tr>
<td>animal</td>
<td>6</td>
<td>206</td>
</tr>
<tr>
<td>aroma</td>
<td>5</td>
<td>61</td>
</tr>
<tr>
<td>asociación</td>
<td>10</td>
<td>49</td>
</tr>
<tr>
<td>causa</td>
<td>5</td>
<td>261</td>
</tr>
<tr>
<td>clase</td>
<td>5</td>
<td>404</td>
</tr>
<tr>
<td>color</td>
<td>5</td>
<td>335</td>
</tr>
<tr>
<td>completo</td>
<td>8</td>
<td>143</td>
</tr>
<tr>
<td>debate</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>decepcion</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td>director</td>
<td>8</td>
<td>173</td>
</tr>
<tr>
<td>escape</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>final</td>
<td>5</td>
<td>417</td>
</tr>
<tr>
<td>frágil</td>
<td>6</td>
<td>44</td>
</tr>
<tr>
<td>ignorant</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>importante</td>
<td>10</td>
<td>348</td>
</tr>
<tr>
<td>imposible</td>
<td>9</td>
<td>275</td>
</tr>
<tr>
<td>informal</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>insertar</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>instancia</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>invitar</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>irregular</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>juvenil</td>
<td>7</td>
<td>38</td>
</tr>
<tr>
<td>liberal</td>
<td>7</td>
<td>44</td>
</tr>
<tr>
<td>local</td>
<td>5</td>
<td>82</td>
</tr>
<tr>
<td>moral</td>
<td>5</td>
<td>342</td>
</tr>
<tr>
<td>natural</td>
<td>7</td>
<td>334</td>
</tr>
<tr>
<td>notable</td>
<td>7</td>
<td>58</td>
</tr>
<tr>
<td>observar</td>
<td>8</td>
<td>82</td>
</tr>
<tr>
<td>personal</td>
<td>8</td>
<td>264</td>
</tr>
<tr>
<td>preparar</td>
<td>8</td>
<td>43</td>
</tr>
<tr>
<td>prudente</td>
<td>8</td>
<td>26</td>
</tr>
<tr>
<td>puro</td>
<td>4</td>
<td>132</td>
</tr>
<tr>
<td>verbal</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>violin</td>
<td>6</td>
<td>17</td>
</tr>
</tbody>
</table>

Average: 6.92  125.5  6.92  123.83  6.78  80.72
<table>
<thead>
<tr>
<th>Homographs</th>
<th>Homograph Controls</th>
<th>English Homographs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target</strong></td>
<td><strong>Length</strong></td>
<td><strong>Frequency</strong></td>
</tr>
<tr>
<td>via</td>
<td>3</td>
<td>104</td>
</tr>
<tr>
<td>lata</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>palo</td>
<td>4</td>
<td>51</td>
</tr>
<tr>
<td>mesa</td>
<td>4</td>
<td>469</td>
</tr>
<tr>
<td>sauce</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>venta</td>
<td>5</td>
<td>42</td>
</tr>
<tr>
<td>playa</td>
<td>5</td>
<td>137</td>
</tr>
<tr>
<td>largo</td>
<td>5</td>
<td>584</td>
</tr>
<tr>
<td>tender</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>sucesso</td>
<td>6</td>
<td>58</td>
</tr>
<tr>
<td>infante</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>atender</td>
<td>7</td>
<td>37</td>
</tr>
<tr>
<td>sensible</td>
<td>8</td>
<td>50</td>
</tr>
<tr>
<td>liar</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>toro</td>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>vaso</td>
<td>4</td>
<td>119</td>
</tr>
<tr>
<td>dice</td>
<td>4</td>
<td>744</td>
</tr>
<tr>
<td>talón</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>surge</td>
<td>5</td>
<td>52</td>
</tr>
<tr>
<td>doble</td>
<td>5</td>
<td>182</td>
</tr>
<tr>
<td>gente</td>
<td>5</td>
<td>597</td>
</tr>
<tr>
<td>quitar</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>placer</td>
<td>6</td>
<td>222</td>
</tr>
<tr>
<td>resumir</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>colegio</td>
<td>7</td>
<td>168</td>
</tr>
<tr>
<td>recordar</td>
<td>8</td>
<td>166</td>
</tr>
<tr>
<td>dame</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>gota</td>
<td>4</td>
<td>47</td>
</tr>
<tr>
<td>pena</td>
<td>4</td>
<td>194</td>
</tr>
<tr>
<td>pulpo</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>horn</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>silla</td>
<td>5</td>
<td>132</td>
</tr>
<tr>
<td>tales</td>
<td>5</td>
<td>252</td>
</tr>
<tr>
<td>pillar</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>collar</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>blanco</td>
<td>6</td>
<td>294</td>
</tr>
<tr>
<td>relieve</td>
<td>7</td>
<td>32</td>
</tr>
<tr>
<td>librería</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td><strong>Average:</strong></td>
<td>5.37</td>
<td><strong>130.18</strong></td>
</tr>
</tbody>
</table>