The Pennsylvania State University
The Graduate School
College of the Liberal Arts

ARE BILINGUALS BETTER LEARNERS?
A NEUROCOGNITIVE INVESTIGATION OF THE BILINGUAL ADVANTAGE

A Dissertation in
Psychology and Language Science

by
Cari Anne Bogulski

© 2013 Cari Anne Bogulski

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2013
The dissertation of Cari Anne Bogulski was reviewed and approved* by the following:

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

Paola E. Dussias  
Associate Professor of Spanish, Linguistics, and Psychology  
Head of the Department of Spanish, Italian and Portuguese

Janet G. Van Hell  
Professor of Psychology and Linguistics and  
Director of the Linguistics Program

Daniel J. Weiss  
Associate Professor of Psychology and Linguistics

Melvin M. Mark  
Department Head of Psychology  
Professor of Psychology

*Signatures are on file in the Graduate School.
Abstract

The general cognitive consequences of bilingualism is a recent topic of intense research interest. A large body of evidence has suggested that bilinguals exhibit several cognitive advantages over matched monolinguals (see Bialystok, Craik, Green, & Gollan, 2009, for a review), including several subdomains of executive function, such as attentional mechanisms, cognitive flexibility, and inhibitory control. Though the extent to which this is true for all previously identified affected aspects of cognition for all types of bilinguals is not yet well known, the precise mechanism underlying such advantages is perhaps even less well understood. In addition, a less well-studied bilingual advantage in foreign language vocabulary learning has been identified for bilinguals, and the origins of this advantage are even less well understood. Perhaps more importantly, the idea that bilinguals may be advantaged at learning more generally has not yet been pursued in the research literature. This dissertation aims not only to better understand the extent and limitations of both of these types of bilingual advantages, but also to investigate how they may or may not be independent from one another, as well as the possibility that bilinguals may have advantages for multiple types of new learning, which may or may not be independent of the advantages in executive function.
Chapter 4: Experiment 1b – Re-examining the Bilingual Advantage in Foreign Language Vocabulary Learning via L2 Translations Using ERPs

4.1 Experiment 1b .................................................................................................................. 91
  4.1.1 Method ...................................................................................................................... 91
    4.1.1.1 Participants ........................................................................................................... 91
    4.1.1.2 Materials and Design ......................................................................................... 93
    4.1.1.3 Procedure ............................................................................................................ 93
    4.1.1.4 EEG acquisition and offline processing ............................................................. 94
  4.1.2 Results: Behavioral Data ......................................................................................... 94
    4.1.2.1 Flanker task ......................................................................................................... 96
    4.1.2.2 Word naming at study: Behavioral data ............................................................ 99
    4.1.2.3 Recognition test task: Behavioral data ............................................................... 101
  4.1.2 Results: ERP Data .................................................................................................. 107
    4.1.2.1 Test task ERP Results ....................................................................................... 107
    4.1.2.1.1. Discussion of Bilingual ERP results ............................................................. 110

Chapter 5: Experiment 2 – Are Bilinguals Advantaged on Tasks of Implicit Learning and Pattern Recognition?

5.1 Experiment 2 ................................................................................................................. 120
  5.1.1 Method .................................................................................................................... 121
    5.1.1.1 Participants ........................................................................................................ 121
    5.1.1.2 Materials and Procedure .................................................................................. 121
      5.1.1.2.1 Statistical Learning Task .............................................................................. 122
      5.1.1.2.2 Operation-Span Task .................................................................................. 125
      5.1.1.2.3 AX-CPT ........................................................................................................ 125
      5.1.1.2.4 Boston Naming Task ................................................................................... 127
      5.1.1.2.5 Dot-Motion Discrimination Task ............................................................... 128
  5.1.2 Results ..................................................................................................................... 128
    5.1.2.1 Statistical Learning Task .................................................................................. 130
    5.1.2.2 AX-CPT ............................................................................................................ 132
    5.1.2.3 Dot-Motion Discrimination Task ..................................................................... 136

Chapter 6: General Discussion

6.1. Summary of Findings ................................................................................................. 140
  6.1.1. Summary of Experiment 1a ................................................................................. 140
  6.1.2. Summary of Experiment 1b ................................................................................ 141
  6.1.3. Summary of Experiment 2 .................................................................................. 142
  6.2 Conclusions ............................................................................................................... 144

References ......................................................................................................................... 156

Appendix ............................................................................................................................. 173
LIST OF FIGURES

Figure 3.1: Types of trials presented in the Flanker task .................................................. 52
Figure 3.2. 32-channel QuikCap electrode montage used for ERP data collection .... 55
Figure 3.3: Study task naming latencies for Experiment 1a ............................................. 64
Figure 3.4: Test task accuracy (d’) for Experiment 1a ....................................................... 66
Figure 3.5a: Test task reaction times on nonword trials for Experiment 1a ............... 67
Figure 3.5b: Test task reaction times on word trials for Experiment 1a ................. 68
Figure 3.6a. Grand average ERP waveforms for monolinguals at the Pre-Test for
   Experiment 1a.............................................................................................................. 71
Figure 3.6b. Grand average ERP waveforms for monolinguals at Test 1 for
   Experiment 1a............................................................................................................. 72
Figure 3.6c. Grand average ERP waveforms for monolinguals at Test 2 for
   Experiment 1a............................................................................................................. 73
Figure 3.6d. Grand average ERP waveforms for monolinguals at Test 3 for
   Experiment 1a............................................................................................................. 74
Figure 3.7. Grand average ERP waveforms for monolinguals from Experiment 1a at
   representative electrode site CZ for all test task presentations .......................... 76
Figure 3.8a. Grand average ERP waveforms for bilinguals at the Pre-Test for
   Experiment 1a............................................................................................................. 80
Figure 3.8b. Grand average ERP waveforms for bilinguals at Test 1 for Experiment
   1a .................................................................................................................................. 81
Figure 3.8c. Grand average ERP waveforms for bilinguals at Test 2 for Experiment
   1a .................................................................................................................................. 82
Figure 3.8d. Grand average ERP waveforms for bilinguals at Test 3 for Experiment
   1a .................................................................................................................................. 83
Figure 3.9. Grand average ERP waveforms for monolinguals and bilinguals from
   Experiment 1a at representative electrode site CZ for all test task
   presentations ................................................................................................................ 85
Figure 4.1: Study task naming latencies for Experiments 1a and 1b ...................... 100
Figure 4.2: Test task accuracy (d’) for Experiments 1a and 1b ..................................................102
Figure 4.3a: Test task reaction times on nonword trials for Experiments 1a and 1b
........................................................................................................................................104
Figure 4.3b: Test task reaction times on word trials for Experiments 1a and 1b ....104
Figure 4.4a. Grand average ERP waveforms for bilinguals at the Pre-Test for
Experiment 1b ..........................................................................................................................109
Figure 4.4b. Grand average ERP waveforms for bilinguals at Test 1 for Experiment
1b ........................................................................................................................................109
Figure 4.4c. Grand average ERP waveforms for bilinguals at Test 2 for Experiment
1b ........................................................................................................................................110
Figure 4.4d. Grand average ERP waveforms for bilinguals at Test 3 for Experiment
1b ........................................................................................................................................111
Figure 4.5. Grand average ERP waveforms for monolinguals and bilinguals from
Experiment 1a and bilinguals from Experiment 1b at representative
electrode site CZ for all test task presentations ..............................................................113
Figure 5.1. The statistical learning task procedure, adapted from Gebhart et al.
(2009) ....................................................................................................................................124
Figure 5.2. The four trial types in the AX-CPT task: (A) An AX trial, (B) An AY trial,
(C) A BX trial, and (D) a BY trial..........................................................................................127
Figure 5.3. Percentage of test items correctly identified for both statistical learning
tasks in Experiment 2 for bilinguals and monolinguals .................................................131
Figure 5.4. AX-CPT reaction time data for monolinguals and bilinguals ............133
Figure 5.5. AX-CPT accuracy data for monolinguals and bilinguals...............134
Figure 5.6a. Accuracy rates across all seven levels of dot-motion coherence for
monolinguals and bilinguals in Experiment 2.................................................................136
Figure 5.6b. Mean reaction times across all seven levels of dot-motion coherence for
monolinguals and bilinguals in Experiment 2.................................................................137
LIST OF TABLES

Table 3.1: Experiment 1a Procedure Overview ................................................................. 48
Table 3.2. Mean age, age of L2 onset, L1 self-ratings, L2 self-ratings, Operation-Span score, and mean naming latencies and accuracy on the Boston Naming test for participants in Experiment 1a................................................................. 59
Table 3.3. Monolingual ERP ANOVA results for Experiment 1a ........................................ 77
Table 3.4. Bilingual ERP ANOVA results for Experiment 1a ................................................. 86
Table 4.1. Mean age, number of days between sessions, age of L2 onset, L1 self-ratings, L2 self-ratings, and Operation-Span score for the monolinguals, bilingual L1 learners, and bilingual L2 learners test in Experiments 1a and 1b ........................................................................................................................................... 95
Table 4.2. Bilingual ERP ANOVA results for Experiment 1b.................................................. 114
Table 5.1. Design of the two counterbalanced languages in Experiment 2...................... 123
Table 5.2. Mean age, age of L2 onset, L1 self-ratings, L2 self-ratings, Operation-Span score, and mean naming latencies and accuracy on the Boston Naming test for participants in Experiment 2 ........................................................................................................ 129
ACKNOWLEDGEMENTS

There are an incredible number of people without whom I would never have completed my degree, and for all of you I am deeply grateful. First, I would like to acknowledge funding from a National Science Foundation Dissertation Improvement Grant (Award #: 1124051), the RGSO Dissertation Support Grant from the College of Liberal Arts at the Pennsylvania State University, the Adele Miccio Travel Award from the Center for Language Science at the Pennsylvania State University, and fellowships from the National Science Foundation, both from Partnerships for International Research and Education and from the Graduate Research Fellowship Program. I’d also like to thank the members of my committee, who have all provided valuable feedback during the process of working through my very complicated dissertation. I have immensely enjoyed our interactions over the past six years, and I hope to continue to stay in touch with each of you in the years to come.

I have had the privilege of working with several dedicated and diligent research assistants, without whom I would almost certainly have taken an extra year to graduate: Will Mainwaring, Nicholas Ray, Eric Jansen, and Sarah Fairchild. Also, to Peiyao Chen, Robert Marker, and Joyce Tam: I feel lucky to have been working in the lab with all of you as managers (though unfortunately never all three at the same time). All of your help with participant running, payment, E-Prime shenanigans, and hardware troubleshooting ("the microphone wasn’t on, Cari, calm
down") has been invaluable. And to Fengyang Ma, who is perhaps the kindest person I have ever met.

I have had the pleasure of working with a brilliant and supportive group of graduate students not just in my lab, but all of the cognitive and CLS graduate students and postdocs B.T.B (before Thomas basement), during the dark, dungeon-y times (Thomas basement era): Colleen Balukas; Grant Berry; Kinsey Bice; Susan Bobb, re-reading your dissertation has helped me write this one!; Matt Brown; Chris Champi; Katie Chapman, you have been my lifeline and dearest fellow 6th-year; Angela Chouinard, thanks for being a great TA!; Carla Contemori; Amelia Dietrich; Katharine Donnelly Adams; Melissa Dracos; Jeremy Fesi, I miss our TV watch parties; Aroline Hanson; Lisa Heimbauer; Alison Hendricks; Nick Henry, thanks for letting me join your trivia team; Kiren Khan, let’s meet up in Toronto some time!; Courtney Johnson-Fowler; Kaitlyn Litcofsky; Jared Linck; Mari-Cruz Martín; Rhonda McClain, I’ll never forget our song(s); Ji Sook Park; Laurne Perrotti; Tyler Phelps, thanks for alerting me to nerd-fighter-dom; Tim Poepsel and your discriminating sensibilities in both research design and musicals; Pablo Requena; Ashley Roccamo; Patricia Román; Patricia Schemp; Hiram Smith; Caitlin Ting; Álvaro Villegaas; Ben Zinszer, I’ve missed having you around since you moved to the palace; and Megan Zirnstein. Additionally, Jason Gullifer, if I had paid you minimum wage for the hours you had helped me through something in the six years we’ve known each other, you could probably retire before you needed to write one of these. Eleonora Rossi, not only did you teach me how to do work with ERPs from the capping to the cleaning to
the .dat files to the analysis to the plotting to the statistics, you’ve been an incredible friend I’ve felt lucky to have gotten to know over the past several years. Jorge Valdés Kroff, I’ve so missed you in the one year since you’ve been gone—your statistical and linguistic consulting, your incredible insight into research design and interpretation, the trips to the Creamery—but I take solace in the fact that you’ve sent me more cat videos since you’ve moved to Philadelphia. And Darren Tanner, thank you so much for all of the help you provided me and the rest of the CLS graduate students and postdocs on ERP methods, interpretation, and EEGLAB/ERPLAB. I probably would have drawn my plots by hand if it hadn’t been for you.

I would also like to thank my impressively patient friends and family for supporting me throughout this entire process. Bill, it goes without saying (though it shouldn’t, so it won’t) that without you, I would have forgotten to eat over the past month, would have run out of clean clothes to wear, would be analyzing my data incorrectly at every turn, and would not have a place to live next month. I’m so lucky that you are my partner. John, thanks for the words of encouragement along the way, sandwiched between all of the entertainment news I’ve fallen behind on. You are my favorite brother. And Mom, you have been such an inspiration to me, and not just since your cancer diagnosis (though have I told you what a wonderful patient you have been to take care of?), but throughout my entire life. Thank you for convincing me to take time for myself. I’m so proud to be your daughter.
And Judy, you’ve been an entire support network to me all by yourself. Over the past six years, I have felt certain at a number of different points that I could not finish. Every single time you convinced me that I should stay. And you were always right. I know with every confidence that you are one of the most exceptional people I have ever met. Thank you for everything, and then everything else.

And for my father. When I graduated from college six years ago, just months before you died, you told me that you were glad that my thesis title was at least something you could understand. Except for the word "neurocognitive," I hope this one would have also done you proud. I miss you.
CHAPTER 1: BACKGROUND

1.1 Introduction.

Despite an overrepresentation of monolingual participants in psycholinguistic research, at least half of the world's language users speak more than one language (Grosjean, 2010), and this figure is almost undoubtedly an underestimate (Grosjean, 2012). Despite what intuition might suggest, an increasing number of studies suggest that bilinguals and multilinguals cannot simply "shut off" one language and "turn on" another. In fact, one of the few assertions in the field of bilingualism that most researchers can agree on is that a bilingual activates both languages in parallel (Costa, Miozzo, & Caramazza, 1999; De Groot, Delmaar, & Lupker, 2000; Dijkstra, Grainger, & van Heuven, 1999; Hoshino & Kroll, 2008; Kroll, Bobb, Misra, & Guo, 2008; Kroll, Bobb, & Wodniecka, 2006; van Heuven, Dijkstra, & Grainger, 1998), at the level of syntax (e.g., Hartsuiker, Pickering, & Veltkamp, 2004), lexical/lemma selection (e.g., Van Hell & Dijkstra, 2002), and phonology (e.g., Marian & Spivey, 2003), even if the intention is only to speak in one language. Yet despite parallel activation of multiple languages, healthy bilinguals typically make very few slips into an unintended language (e.g., Gollan, Sandoval, & Salmon, 2011; though see Abutalebi, Della Rosa, Tettamanti, Green, & Cappa, 2009 for one example of a breakdown in this control in a bilingual aphasic patient). Although there is still debate about the nature of cross-language activation (e.g., Assche, Duyck, & Hartsuiker, 2012; Finkbeiner, Almeida, Janssen, & Caramazza, 2006), there is
agreement that there must be mechanisms that regulate the co-activation of two (or more) languages realized in the bilingual brain.

Given such parallel activity of multiple languages, it is a reasonable hypothesis that this type of mental juggling that a bilingual must employ to successfully select targets from the appropriate language for a given context will have consequences for general cognition and the brain. In fact, an increasingly large body of research has been devoted to investigating the possibility that, for this reason, bilinguals may outperform monolinguals on a variety of general cognitive tasks (see Bialystok, Craik, Green, & Gollan, 2009 for a recent review). Rather than viewing bilingual speakers as non-representative of typical language users, this hypothesis emphasizes the role of the bilingual as the expert language user, providing invaluable insight into general cognitive and linguistic architectures and their interactions. The hypothesis that bilinguals may exhibit certain cognitive benefits relative to monolinguals is one that addresses fundamental issues about the interconnectivity of cognitive and linguistic systems, and is not simply an investigative examination of atypical populations.

In addition to the widespread evidence suggesting general cognitive benefits for bilinguals, a comparatively small literature has suggested that bilinguals may be more adept than their monolingual peers at learning vocabulary in a foreign language (Kaushanskaya & Marian, 2009a; Keshavarz & Astaneh, 2004; Papagno & Vallar, 1995; Sanz, 2000; Van Hell & Mahn, 1997). However, very few studies have addressed possible connections between the two. A few studies have examined the
role of phonological working memory underlying this particular bilingual advantage, which seems to be at least partially, if not entirely, independent of the advantage in foreign language vocabulary learning (Kaushanskaya & Marian, 2009a, 2009b; Kaushanskaya, 2012). Hence, the relationship between bilingual benefits to foreign language vocabulary learning and bilingual benefits to executive function is not fully understood.

The goal of the present dissertation is to investigate not only the possibility that bilinguals may have a much broader advantage in learning generally, but also to ask whether such an advantage in learning is related to the bilingual advantages reported for executive function and other domain-general cognitive benefits. By testing both monolinguals and bilinguals on a series of linguistic and cognitive tasks using behavioral and neuroscientific methods, this dissertation aims to investigate the mechanism underlying several of the reported bilingual advantages and the relationship of such benefits to one another. This research has potentially profound consequences not just for bilingual models of language processing and the brain, but for all language users as well as for models of cognition, adding to our understanding of the interactions between how language experience may be related to these various types of benefits, including executive function but also new learning. That is, bilingualism as a phenomenon may provide an especially useful lens through which to investigate how language and cognitive processing interact, which may uncover new routes to facilitate learning and general cognition.

1.2 Bilingual advantages in executive function.
One realm in which differences in cognition between monolinguals and bilinguals have been reported is that of executive function (e.g., Bialystok et al., 2009). An umbrella term for several related but independent cognitive mechanisms, executive function, according to one widely-cited interpretation (Garon, Bryson, & Smith, 2008; Miyake et al., 2000), emphasizes an overarching attentional mechanism, reigning over three distinct, yet related components of executive function: mental set shifting (or cognitive flexibility), updating (or working memory), and response inhibition. The "distinct, yet related" aspect of these three components, in addition to the overarching attentional mechanism, leaves open the possibility for multiple interactions. Hence, when one group may outperform another on a particular task thought to require executive function, the extent to which each mechanism/component is involved is highly debatable.

For example, bilinguals have frequently been reported to exhibit an advantage on tasks thought to require inhibitory control, relative to monolinguals (Bialystok & Martin, 2004; Bialystok & Viswanathan, 2009). It has been argued that this benefit to bilinguals stems from the need to suppress activated linguistic targets not just within a particular language (as all monolinguals do), but also across languages, in order to avoid slipping into an unintended language (e.g., Kroll et al., 2008). In fact, bilinguals of all ages rarely make such cross-language errors (see Gollan, Sandoval, & Salmon, 2011 for evidence from older bilingual adults, and Petitto et al., 2001 for evidence from infants learning to differentiate between multiple languages in the input). It is therefore hypothesized that bilinguals must
make extensive use of a domain-general control mechanism to effectively negotiate the parallel activation of multiple languages, and an efficient inhibitory control mechanism has been proposed to underlie this process. Specifically, it is a suppression of interference that has been reported to be enhanced in bilinguals relative to monolinguals, rather than suppression of a response (Luk, Anderson, Craik, Grady, & Bialystok, 2010).

Within the executive function model proposed by Miyake et al. (2000), however, it is difficult to determine which mechanism and/or component(s) underlie inhibitory control. Inhibitory control may, in fact, be related to cognitive flexibility (shifting between targets), response suppression (inhibiting a response to a target), and/or attentional control (identifying the activated targets). The involvement of each of these particular aspects of executive function may also differ among the tasks purported to involve inhibitory control.

However, in addition to having an advantage over monolinguals in inhibitory control, bilinguals have also been argued to have advantages on other aspects of executive function, such as cognitive flexibility. Prior and MacWhinney (2010) administered a task-switching task to young adult bilingual and monolingual participants and found a bilingual advantage, specifically manifested in a reduced switching cost (i.e., experiencing less interference when switching from responding to one aspect of a stimulus, such as color, to another, such as shape, during a mixed block). Prior and MacWhinney argue that a reduction in switching costs may be related to a reduction in proactive interference, and further suggest that the
advantage demonstrated by these early bilingual participants may actually be one of enhanced attentional control processes.

It makes intuitive sense that bilinguals might have an advantage on tasks that require switching between sets, due to the preponderance of evidence that bilinguals are constantly juggling activated candidates in multiple languages. Code-switching—or the fluent, deliberate, and systematic use of multiple languages within a sentence or discourse—occurs naturally in many bilingual communities worldwide (e.g., Poplack, 1980), suggesting that language switching is a particular skill that is honed in at least some bilingual individuals. However, experimental evidence from language-switching tasks suggests that even among proficient bilinguals, language-switching has a cost. Laboratory-based language-switching paradigms differ from naturally-occurring code-switching in that switches are determined by the experimenter and/or the experimental paradigm, and most often lack the syntactic and lexical cues that govern code-switching. Despite sociolinguistic evidence of code-switching in language production (and, of course, the reality that code-switched speech must also be comprehended), the most frequent finding in the language-switching literature is asymmetrical switch costs: switching between languages is more costly than non-switching, but switching from the weaker second language (L2) into the more dominant native language (L1) is more costly than the reverse (e.g., Meuter & Allport, 1999). Again, most of this evidence comes from paradigms in which a switch in language was either unpredictable and/or devoid of linguistic context that might otherwise cue the
comprehender that a switch is about to occur. It has been argued (e.g., Meuter & Allport, 1999) that this asymmetry is caused by a greater need to inhibit the stronger and more automatically activated L1 when using the weaker L2, as the L1 is highly likely to be activated and must be successfully inhibited in order to speak the L2. Hence, it is thought that this successful inhibitory process has a cost when switching back into the L1, as the inhibition employed during L2 production must now be overcome. Indeed, among balanced bilinguals (i.e., those who have approximately equal proficiency in both languages), this asymmetry may not be present at all (Costa, Santesteban, & Ivanova, 2006).

More recent psycholinguistic evidence suggests that contextual effects may be critical in determining whether language-switching costs are present or absent. For instance, when lexical language-switching (i.e., single words) is under the control of the bilinguals themselves, costs to language production vanish (Gollan & Ferreira, 2009). Additionally, evidence from sentence-level language-switching (i.e., comparing blocks of code-switched sentences to unilingual sentences presented in blocks) suggests that code-switched sentences (like those that are naturally produced by bilinguals in code-switching communities) introduce no additional processing costs to the comprehension of words in those sentences (Gullifer, Kroll, & Dussias, 2013). However, the degree to which these effects are due to the experience of knowing multiple languages or to the experience of code-switching is yet unknown.
In addition to inhibitory control and cognitive flexibility, bilingual benefits have been reported for a variety of other constructs that fall under the umbrella of executive function: attentional control (Martin-Rhee & Bialystok, 2008), conflict resolution (Bialystok, 2010; Carlson & Meltzoff, 2008; Costa, Hernández, & Sebastián-Gallés, 2008), selection of goal-relevant information (Colzato et al., 2008), working memory (Kroll, Michael, Tokowicz, & Dufour, 2002), interference suppression (Martin-Rhee & Bialystok, 2008), reduction of proactive interference (Bialystok & Feng, 2009), and monitoring (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009). It is unlikely that all of these reported advantages are independent of one another, but it remains possible that there are multiple cognitive benefits for bilinguals. Moreover, discussing "bilingual advantages" is a bit misleading, as bilingualism includes a highly diverse set of language experiences, ranging from individuals who are highly proficient and balanced across languages, to heritage speakers who have been educated in a language different from their home language, to early learners of a second language, and beyond. To assume that individuals across the complex spectrum of bilingualism—who spend different amounts of time speaking each language, have different proficiencies across languages, switch between languages with varying frequency, etc.—would be exercising linguistic and cognitive mechanisms to precisely the same degree seems unlikely.

Green and Abutalebi (2013) expand on this with their adaptive control hypothesis, which proposes that bilingualism—as well as other types of expertise—
differentially exercise a control network depending on particular language experience. Hence, Green and Abutalebi argue that the types of executive function advantages that bilinguals exhibit will vary depending on their particular bilingual experience. Determining the precise role of different aspects of bilingualism—such as proficiency, age of acquisition, and frequency of code-switching, to name a few—is resource-intensive, and would involve testing an enormous variety of bilingual speakers across the world. And at this point in the landscape of the cognitive consequences of bilingualism, these questions as yet remain unanswered.

1.3 A bilingual advantage in foreign language vocabulary learning.

In addition to the advantage that has been reported for executive function, bilinguals have also been shown to outperform their monolingual peers in learning vocabulary words in an unfamiliar language (Bogulski & Kroll, under review; Kaushanskaya & Marian, 2009a, 2009b; Kaushanskaya, 2012; Keshavarz & Astaneh, 2004; Papagno & Vallar, 1995; Sanz, 2000; Van Hell & Mahn, 1997). However, this effect has received very little experimental attention, and having only a few studies reporting it makes it difficult to determine the nature of this particular advantage. Perhaps it is because this result is such an intuitive one: learning a second language seems to help when learning a third. Yet because this phenomenon has been under-investigated, it is yet unknown what mechanisms underlie such an advantage, even if it is an intuitive one.

One difficulty in identifying the source of the bilingual advantage in foreign language vocabulary learning is the heterogeneity of the small number of studies
that have been done. Papagno and Vallar (1995) reported this advantage, and in their study, they taught polyglots (operationally defined as individuals who spoke at least three languages) and non-polyglots a set of only 8 transliterated Russian words into Italian (a language all participants were described as knowing fluently), and immediately tested their memory for these words via cued recall. Ten iterations of the study-test procedure were conducted for each participant. The researchers found an advantage in learning these quasi-Russian translations of Italian words for the polyglot learners over the non-polyglot learners: the polyglot learners were able to remember more of the transliterated Russian words in the cued recall paradigm at each of the ten iterations of the test than were the non-polyglot learners. Quite remarkably, the polyglot learners did not demonstrate an advantage in paired associate word learning (in Italian), which was taught and tested through the same method as the transliterated Russian words. This suggests that it is not that polyglots have better memories in general than do non-polyglots, but rather that there is something specific about a polyglot experience that aids the acquisition of foreign language vocabulary.

Despite the compelling evidence presented in the Papagno and Vallar (1995) study, however, the experiment had a number of features that may have contributed to the results. For instance, only ten learners per group were tested, and among each already very small group there was quite a degree of variability; among the polyglot participants, five spoke four languages and five of them spoke three. Perhaps even more puzzling is that the non-polyglot learners were, in fact, bilingual
themselves, having learned a second language (but not a third or a fourth) in school. Not only were the non-polyglots bilingual, but they had learned their L2 beginning on average at the age of 11, which is the same age reported for polyglots’ age of acquisition of their first foreign language. Moreover, none of the subjects had multilingual parents. Hence, the advantage reported for the polyglots in this study could potentially be one of self-selection: the students who excelled at learning an L2 at an early age went on to study an L3 and possibly an L4, and did better than those who did not at learning the unfamiliar vocabulary presented in this experiment.

A more direct comparison of bilingual and monolingual performance (as opposed to polyglot and non-polyglot) was conducted by Van Hell and Mahn (1997). Though their study was designed to compare two types of foreign language vocabulary learning methods, they taught and tested one group of Dutch-English bilinguals and one group of functionally-monolingual English speakers. The Dutch-English bilingual participants all had at least 6 years of classroom instruction in English (i.e., they had learned their first foreign language after early childhood), and at least 3 years of French and at least 2 years of German. The English monolinguals had little to no experience in foreign language learning. The bilingual participants were taught and tested on a set of 60 Spanish words learned via Dutch (L1) translations, and the monolingual participants were taught and tested on a set of 56 Dutch words learned via English (L1) translations. Both groups were administered three tests: an immediate test, a test after a delay of one week, and a test after a
delay of two weeks. Even with slightly more items to learn, the Dutch-English bilinguals outperformed the English monolinguals, regardless of learning method or test delay (immediate, one-week delay, or two-week delay).

The results of this study further support the idea that second language classroom experience predicts better foreign language vocabulary learning. However, due to the nature of the multilingualism of the participants tested in the studies mentioned thus far, the results from neither Papagno and Vallar (1995) nor Van Hell and Mahn (1997) can rule out the possibility that it is only with exposure to multiple foreign languages that a benefit emerges. This is critical for understanding what possible mechanism(s) may underlie this advantage in foreign language vocabulary learning.

In a departure from the late-learner, classroom-based-L2 multilinguals previously tested, Kaushanskaya and Marian (2009a) tested two groups of early bilinguals—an English-Spanish group and an English-Chinese group—along with a group of English monolinguals. Participants in the two bilingual groups reported their L2 usage as primarily family-based (as opposed to the classroom), and their age of acquisition as being quite early in development ($M = 5.44$ years for the English-Spanish bilinguals; mean = 2.21 years for the English-Chinese bilinguals). All participants were taught 48 words in an artificial language via English (L1) translations. Kaushanskaya and Marian (2009a) found further evidence in support of a bilingual advantage in foreign language vocabulary learning, for both recall and recognition tests, and for both immediate and one-week-delayed tests. Additionally,
the three groups were given a digit-span task to index phonological working memory, which was not significantly different between the groups. A related study using the same artificial language (Kaushanskaya & Marian, 2009b) testing only monolinguals and English-Spanish bilinguals used two modes of learning input: unimodal (hearing-only) and bimodal (hearing and seeing). Importantly, in the bimodal input case the letter-to-phoneme mappings conflicted between the artificial language and English or Spanish, introducing potential interference. In this study, the authors again found evidence for a bilingual advantage in foreign language vocabulary learning. Moreover, the bilingual learners experienced less interference in the bimodal input case than did the monolingual learners. The monolingual learners were better able to learn the new vocabulary via unimodal input, whereas the type of input (unimodal or bimodal) did not affect learning for the English-Spanish bilinguals.

A more recent set of studies reported by Kaushanskaya (2012) investigated possible limitations of the bilingual advantage in foreign language vocabulary learning by asking English-Spanish bilinguals and English monolinguals across two experiments to learn foreign language vocabulary that was either phonologically similar or dissimilar to languages they already knew. In addition, the monolinguals were divided into subgroups based on their phonological working memory scores (high or low). For learning of foreign language vocabulary with familiar phonology, bilinguals outperformed both monolingual groups (high- and low-phonological working memory span), who did not differ from one another. However, for the
foreign language vocabulary with unfamiliar phonology, while the bilinguals again outperformed both groups of monolinguals, the high-span monolinguals exhibited a slight benefit over the low-span monolinguals. These results suggest that, while phonological working memory may be related to some facility in foreign language vocabulary learning, it cannot fully account for it. Recent evidence from Bogulski and Kroll (under review) further supported the limited role of working memory in benefits to foreign language vocabulary learning. While monolinguals with high working memory spans were slightly more "bilingual-like" in their performance—meaning slower performance during the study task, but more accurate performance during a test task—the bilingual group still outperformed even the high-span monolinguals.

A few additional studies have found support for the presence of a bilingual advantage in foreign language vocabulary learning. Sanz (2000) found that adolescent Spanish-Catalan bilinguals outperformed Spanish monolinguals on a standardized 75-item multiple choice assessment of English structure and vocabulary, even though both groups had received similar education in English. Similarly, Keshavarz and Astaneh (2004) found that both Turkish-Persian and Armenian-Persian bilingual learners of English outperformed Persian monolingual learners of English on a cued-recall task consisting of 36 items embedded in a meaningful sentence context. However, the Armenian-Persian bilinguals slightly outperformed the Turkish-Persian bilinguals, which is notable due to the nature of the two groups’ bilingualism: the Armenian-Persian bilinguals were educated in
both Armenian and Persian, and the Turkish-Persian bilinguals were heritage
speakers of Turkish. Perhaps what is most striking about this result is that two
bilingual groups presenting very different types of bilingualism both outperformed a
monolingual group, suggesting some generalizability of the bilingual advantage in
foreign language vocabulary learning.

1.3.1 Extending the advantage to learning via L2 translations. Taken
together, these previous studies suggest that the bilingual advantage in foreign
language vocabulary learning is not restricted to bilinguals who speak only
particular language combinations learning only vocabulary in certain languages.
One as yet unexplored possibility, however, that remains is that this particular
advantage could be limited to bilinguals learning L3 vocabulary via L1 (or dominant
language) translations. This is an important distinction, as an advantage restricted
to only bilinguals learning L3 vocabulary via L1 translations would suggest a
fundamentally different mechanism than if bilinguals demonstrated the advantage
learning either via L1 or L2 translations. Specifically, if the bilingual advantage in
foreign language vocabulary learning is not limited to learning via L1 translations,
this would suggest that a very general cognitive mechanism may be underlying the
effect. Such a mechanism may be rooted in memory, speed of processing, attention,
or any other language-independent systems. Alternatively, if the bilingual advantage
were restricted to cases of bilinguals learning L3 vocabulary only via L1 translations
(the only route thus investigated), this could suggest that a more specific
mechanism was involved. It is also possible that, in the case of the advantage being
restricted to learning via L1 translations, a general mechanism is involved, but the context of learning via L1 may place entirely different demands on the cognitive system than learning via L2. Among the studies previously reviewed, these two possibilities cannot be teased apart.

Bogulski and Kroll (under review) aimed to investigate the generalizability of the bilingual advantage by teaching foreign language vocabulary words to three groups of bilinguals and one group of monolinguals. All groups learned Dutch vocabulary (unfamiliar to all participants) via English translations. Critically, English was the native language (L1) for the monolinguals and for one group of bilinguals (English-Spanish), but was the L2 for the other two groups of bilinguals (Spanish-English and Chinese-English). Bogulski and Kroll found an advantage for only the bilinguals who learned Dutch vocabulary via their dominant language; that is, for the bilinguals whose dominant language was English. They argued that this specificity of the bilingual advantage in foreign language vocabulary learning was due to increased interference from English that the English-dominant bilinguals experienced during the study task. The English-dominant bilinguals were much slower than the other participants (both monolinguals and non-English-dominant bilinguals) to name the English translations of the Dutch words presented during the study task. These increased naming latencies were interpreted as an index of greater interference, which ultimately led to better retention of the foreign language vocabulary, at both immediate and later recognition tests.
The role of interference, or even difficulty, encountered during learning has long been known to have an impact on later retention. Schmidt and Bjork (1992) argued in a review of both language and motor learning tasks that increased difficulty during the training/learning phases ultimately leads to better retention for both motor tasks (difficulty operationalized as blocked vs. randomized presentation of a motor task to be learned) and language tasks (difficulty operationalized as the duration of lag between training and test of novel names). A more recent foreign language vocabulary learning study also manipulated the amount of time between study and test of 200 words in various Inuit dialects and their English translations (Pashler, Zarow, & Triplett, 2003). Pashler et al. found that, while initially a short lag was helpful for learners (i.e., short lags led to the highest retention), a delayed test revealed that the detriment to long-lag trials relative to short-lag trials was completely overcome. In short, increased difficulty during study is initially costly, but ultimately can be beneficial. Similarly, experimental work with aphasic and other language-impaired populations has suggested that errorless-learning—a technique that emphasizes error avoidance in production during re-learning—is ultimately detrimental for patients (see Middleton & Schwartz, 2012 for a recent critical review), arguing that such practice minimizes difficulty during training, which leads to poorer patient outcomes.

One hypothesis that may be proposed based on the results reviewed is that the bilingual advantage in foreign language vocabulary learning is not an advantage in learning per se, but rather a direct outcome from experiencing greater
interference during study than monolinguals and bilinguals learning via L2 translations. As previously mentioned, there is robust evidence for increased inhibitory processes for bilinguals during L2 processing relative to L1 processing, as during L2 processing, the dominant L1 must be constantly suppressed. Hence, during a foreign language vocabulary learning task where a bilingual is learning vocabulary via L1 translations, the L1 must continually be inhibited in order to attend to, encode, and remember the new vocabulary. This task would then be made especially difficult if the participants were additionally asked to produce the L1 target aloud on each study task trial, as was the case for the participants tested by Bogulski and Kroll (under review). This notion aligns with the recent adaptive control hypothesis from Green and Abutalebi (2013), which proposes that bilingualism tunes the cognitive control system inherent in all language users differently from monolingualism, and that different types of bilingualism will tune this system slightly differently. This hypothesis is consistent with the idea that bilinguals and monolinguals are essentially performing the same task utilizing the same cognitive architectures, but a different cognitive experience would be induced in the bilinguals learning via L1 translations.

Some problematic evidence for this hypothesis was reported by Kaushanskaya and Marian (2009b), who frame their evidence for the bilingual advantage in foreign language vocabulary learning as a reduction of experienced interference. By teaching via two methods—one with increasing interference from incongruent phonology-orthography, one without—the researchers found that,
although the bilinguals outperformed the monolinguals overall, this advantage was greatest for the words learned via the method with the greatest interference. While this may seem consistent with the adaptive control hypothesis, it would not explain why the monolinguals, who were clearly affected by the interference manipulation, learned the words acquired via the more difficult method less well than words acquired via the easier one. Additionally, bilinguals performed equally well for words taught via both methods, further arguing against increased difficulty during training ultimately leading to better retention. As study-task performance was not assessed in this particular experiment, the degree of interference experienced during the training portion of the study cannot be determined. But given the outcomes and the literature previously reviewed, a possible explanation for this result is that the bilinguals were unaffected by the manipulation because they were already experiencing a high level of interference in both conditions, as they were learning via L1 translations. Additionally, if the cognitive consequences of bilingualism are viewed as a combination of costs and benefits in the manner of Pashler et al. (2003), for monolinguals, there is an initial cost (bimodal > unimodal input) for learning via difficult conditions, but this cost is ultimately negated at delayed tests: there is no significant effect of learning method for the delayed recognition test. Furthermore, while a cost is still present for the production test, it is diminished.

1.3.2 Using ERPs to assess vocabulary learning. To date, the bilingual advantage in foreign language vocabulary learning has been examined
experimentally using primarily exclusively behavioral methods. However, event-related potentials (ERPs), which are averaged waveforms of an electroencephalogram (EEG) of the brain, time-locked to critical experimental events and recorded at the scalp, have been successfully used in studying complex language processing tasks, including foreign language vocabulary learning more generally (i.e., not only among bilinguals). Over many years and after many replications, particular patterns of brain activity emerge after specific language processing events. Relevant to the study of foreign language vocabulary learning is the N400 component. Named for its polarity (negative) and the latency of its approximate peak amplitude post stimulus onset (400 milliseconds), the N400 component is observed most clearly at centro-parietal sites, and emerges in language processing tasks when a stimulus is difficult to integrate into a meaningful context. For instance, it was first documented by Kutas and Hillyard (1980), who found an N400 component in response to sentences such as "I take coffee with cream and *dog*" (with brain responses time-locked to the presentation of the word *dog*), relative to sentences such as "I take coffee with cream and *sugar* (with brain responses time-locked to the presentation of the word *sugar*). A later study (Kutas & Hillyard, 1984) not only replicated this effect, but also showed that the least predictable words (i.e., low cloze probability) elicited the largest N400, highly predicted words elicited the smallest (or none at all), and words with medium-level cloze probability fell between the two.
These seminal studies contrasted with previous ERP research on the P300 (a positive-going waveform, peaking around 300 milliseconds post stimulus onset), which had been previously reported in response to stimuli that were surprising in some way, such as in the oddball paradigm (e.g., Sutton, Zubin, & John, 1965). In the oddball paradigm, participants attend to (auditorily or visually) rapidly presented stimuli that are overwhelming of one type (e.g., 95%), but will occasionally experience stimuli of a different type (e.g., 5%). A P300 component is elicited for these "oddball" stimuli, relative to the brain responses for the more common stimuli. Hence, one might have expected that upon reading the word *dog*, which is a surprising event, a P300 would be evoked. However, the presence of the N400—a waveform that not only differs in polarity and latency of peak amplitude, but is also morphologically distinct—was, in and of itself, a surprising event.

The N400 is not a component that is elicited for all surprising linguistic events, as oddball paradigms using single words elicit P300 components (Kutas, McCarthy, & Donchin, 1977; Shelburne, 1973). Instead, the N400 seems to be specifically elicited in cases where the expectation in a coherent context is violated (see Kutas & Federmeier, 2011, for a recent review of the N400 across a wide variety of experimental domains). In fact, the N400 has been observed in such disparate tasks as anomalous pictures in sentence context (Nigam, Hoffman, & Simons, 1992); incongruent trials on a Stroop task (Rebai, Bernard, & Lannou, 1997); and word/music pairs, defined a priori as incongruous (Koelsch et al., 2004). Perhaps surprisingly, the N400 has been elicited in response to pseudowords
relative to real words in single word presentation (e.g., Bentin, McCarthy, & Wood, 1985; Curran, 1999; Rugg, 1983). Hence, it may be more accurate to say that the N400 reflects a violation of contextual or meaningful expectation.

In recent years, the N400 has been used to investigate language learning across a wide range of proficiency in participants. For example, Osterhout, McLaughlin, Pitkänen, Frenck-Mestre, and Molinaro (2006) investigated the brain responses to grammatical and ungrammatical French constructions of native-French speakers and college-aged L2 learners of French after the latter had 1, 4, and 8 months of French classroom instruction. For both types of grammatical violations tested (subject/verb agreement and number agreement), native-French speakers showed robust P600 responses—a positive-going waveform that peaks approximately 600 milliseconds after the stimulus onset, and is typically elicited to violations of syntactic expectation. The highest performing French-learners, on the other hand, showed an N400-like response to these violations at 1 month of instruction, an emerging P600-like response at 4 months, and a robust and near-native-like P600 at 8 months. The authors argue that ERPs—and the N400 and P600 components in particular—can be examined and compared in native speakers and L2 learners to examine the native-likeness of their language processing.

What is most intriguing about the use of ERPs in L2 processing for the purposes of foreign language vocabulary learning is their potential to reveal evidence of learning in the absence of behavioral evidence. McLaughlin, Osterhout, and Kim (2004) did exactly this, comparing a group of college-aged L2 learners of
French a matched group of non-French-learners. These participants were given a series of prime-target pairs and asked to make a lexical decision judgment for each target. The primes consisted of real French words, and the targets were either related French words, unrelated French words, or French-like pseudowords. Brain responses were time-locked to the presentation of each of these types of targets. The behavioral responses were unremarkable: after approximately 14 hours of instruction, the French learners were at chance in their ability to detect real French words from French-like pseudowords, as were the non-learners. After approximately 63 hours of instruction, the learners’ ability to discriminate real French words had improved, and by approximately 138 hours of instruction, this ability had improved even more. Unsurprisingly, the non-learners never exceeded chance levels of detection. However, the remarkable result was found in the comparison of the ERP and behavioral data. Although the N400 component emerged at the second and third tests for both the French-like pseudowords (where it was largest) and the unrelated French words (where it fell between the related French words and the French-like pseudowords), the N400 also emerged at the first test only for the pseudowords, despite that the learners had been unable to distinguish real French words from pseudowords in their behavioral responses.

What studies such as McLaughlin et al. (2004) and others (e.g., Tokowicz and MacWhinney, 2005, who used ERPs to evaluate L2 learning of syntactic structures and found evidence for learning in the ERP record, but not in the behavioral measures), have demonstrated is that ERP methodology can be used as a tool
alongside behavior methods to further examine foreign language vocabulary learning. Behavioral measures such as reaction times and overall accuracy, while useful and certainly not to be overlooked, are limited in their ability to assess learning as it is happening. That is, by the time a participant is able to correctly identify a word in a foreign language, the learning process itself is complete. ERPs, as utilized by McLaughlin et al. (2004) and Tokowicz and MacWhinney (2005), on the other hand, have the potential to reveal language learning as it is taking place. Or, at the very least, using ERPs can potentially offer an alternative to viewing foreign language vocabulary learning as a binary outcome (either learned or not learned). In conjunction with behavioral data, the use of ERPs allows for a much more nuanced view of language learning processing, particularly during early phases of learning.

1.4 Enhanced probabilistic inference (i.e., statistical learning) in bilinguals?

A question raised by the presence of a bilingual advantage in foreign language vocabulary learning is whether bilingual learners are simply just better learners, more generally. While this possibility has not been specifically raised with regard to the advantage in foreign language vocabulary learning, one avenue of research related to this idea is statistical learning.

Statistical learning was first developed as a theoretical explanation and methodology that avoids positing innate linguistic knowledge for how infants are able to accomplish one of the first steps in the process of language learning: extracting meaningful segments from a continuous speech stream (e.g., Saffran,
Aslin, & Newport, 1996). This is not a trivial problem, as anyone who has overhead an unfamiliar language knows, as within the acoustic signal there are no overt cues to word boundaries. However, while there are no overt cues in the speech signal, there are regularities that can be extracted. Consider the phrase "pretty baby." The syllable /ˈpri/ is fairly uncommon in English, and is frequently followed by /ti/, as one of the most common words in English containing the syllable /ˈpri/ is within the word "pretty." On the other hand, /ˈbeɪɪ/ is fairly common, and can be found in words such as "basement", "bacon," and "E-bay." For every given pair of syllables in a language, some probability exists that one will follow the other, often labeled a transitional probability. Even when considering infant-directed speech, this transitional probability is very high between /ˈpri/ and /ti/, and slightly lower for /ˈbeɪɪ/ and /bi/. The probability that /ˈbeɪɪ/ will follow /ti/, however, is quite low. For, as common as the phrase may be in infant-directed speech, these two words appear apart from one another more often: "pretty flower", "pretty house", "pretty kitty", "Aren't you pretty?", "baby names", "tiny baby", "baby dinosaur", etc. The ability to attend to and learn these transitional probabilities is referred to as statistical learning, and is thought to underlie the implicit language learning all language-exposed infants experience.

Statistical learning has been observed not only in infants (e.g., Saffran, Aslin, & Newport, 1996; Saffran, 2001), but also in adults (e.g., Saffran, Johnson, Aslin, & Newport, 1999; Saffran, Newport, & Aslin, 1996), suggesting that there is no critical period for this ability. Moreover, research has further suggested that this
mechanism is a domain-general one that operates outside of the auditory domain and can be observed in the visual domain as well (Fiser & Aslin, 2001, 2002; Kirkham, Slemmer, & Johnson, 2002; Saffran et al., 1999), although the extent to which learning is equally likely to occur across multiple modalities (including tactile) has been called into question (Conway & Christiansen, 2005).

Some recent evidence has suggested that bilingual children have no advantage on a statistical learning task relative to monolinguals children (Yim & Rudoy, 2013). Another recent study found no advantage for bilingual adults relative to monolinguals on an implicit-learning task of an artificial grammar (Grey, 2013). As statistical learning is an implicit process in which participants often have no awareness that they have learned at all, the comparison to uninstructed grammar learning contexts seems appropriately analogous. However, other studies have found evidence for a bilingual advantage in statistical learning for adult bilinguals using a Morse code language (Bartolotti, Marian, Schroeder, & Shook, 2011). Hence, the effect of bilingual language experience on statistical learning is not yet clear, and requires further investigation. Evidence outside of bilingualism has, however, suggested that other types of language knowledge may directly impact the statistical learning ability. Evans, Saffran, and Robe-Torres (2009) found that children with specific-language impairment were unable to perform a statistical learning task as well as typically-developing children, and Conway, Baurnschmidt, Huang, and Pisoni (2010) found that sensitivity to sequential structure predicted better statistical learning, after controlling for a variety of additional cognitive factors: working
memory, intelligence, attentional control, inhibitory control, and vocabulary knowledge. These results suggest that particular types of experience can impact the ability to learn these statistics, but the specific role of bilingualism—which, as previously mentioned, is not a uniform experience—is just beginning to be investigated as one such predictor.

Due to the potential domain-generality of the statistical learning mechanism, as well as the potential domain-general cognitive consequences already reviewed, one hypothesis regarding the relationship between bilingual advantages in executive function and bilingual advantages in foreign language vocabulary learning is that the latter may be related to an enhanced ability to extract patterns from noisy input. On the surface, it may seem like foreign language vocabulary learning and statistical learning are completely disparate tasks, and that skilled performance on one is unlikely to predict skilled performance on the other. However, some recent fMRI evidence suggests that the ability to extract such statistics from an input in an unfamiliar language supports the idea that this ability may vary among individuals. Veroude, Norris, Shumskaya, Gullberg, and Indefrey (2010) asked Dutch speakers to listen to a weather report in Chinese, which was a language the participants were unfamiliar with. Then, the participants were given an auditory word recognition task that contained words that had been present in the weather report. The brain responses of the participants who performed best on this task as measured by this test were then compared to those who did less well. The fMRI results demonstrated that the better learners exhibited stronger functional connectivity relative to the
non-learners in brain regions associated with phonological rehearsal (the left supplementary motor area, the left precentral gyrus, the left insula, and the left rolandic operculum). Hence, expertise in statistical learning not only varies from one individual to another, but this variation has a neurophysiological reality. In the context of monolingual and bilingual comparisons, this result raises the possibility that bilinguals might be one such type of expert, and in particular, that bilingual brains may process linguistic stimuli from an unfamiliar language differently than monolinguals.

Some additional evidence that bilingual experience may impact statistical learning comes from research with infants. One of the most striking results from this literature comes from a series of experiments conducted by Kovács and Mehler (2009), who tested bilingual and monolingual 7-month-old infants on a test involving anticipation of a visual reward on a computer screen. Infants at seven months have typically not yet begun to speak, so infants in this study were selected based on exposure to monolingual or bilingual parental input. Critically, the experiment was designed such that the particular pattern (visual and auditory stimuli were used in different experiments) that predicted the visual reward for the infants switched after nine trials, such that a new pattern had to be learned in order to correctly anticipate the new location of the visual reward. While both monolingual and bilingual infants were able to anticipate the location of the visual reward over the first nine trials in all three experiments, only the bilingual infants
were able to re-learn the new set of cues and correctly anticipate the new location of the visual reward in the latter nine trials.

While the Kovács and Mehler (2009) study found evidence of cognitive advantages for bilingual infants, other studies report results that might be interpreted as disadvantages for bilingual infants relative to monolinguals. Bilingual infants are less able to utilize disambiguation—a word-learning heuristic—in a preferential-looking task that presented a familiar and an unfamiliar object while a novel word was heard (Byers-Heinlein & Werker, 2009). Bilingual infants are also slower to attune to phoneme contrasts in their native language(s) (Bosch & Sebastián-Gallés, 1997; Navarra, Sebastián-Gallés, & Soto-Faraco, 2005; Sebastián-Gallés & Bosch, 2009). These differences in bilinguals seem adaptive, however, given the additional challenges that bilingual language input poses for an infant learner. The idea that the tuning of linguistic and cognitive networks may have benefits (e.g., Kovács & Mehler, 2009) as well as other, potentially less-advantageous consequences (e.g., Bosch & Sebastián-Gallés, 2003; Byers-Heinlein & Werker, 2009), is consistent with Green and Abutalebi’s (2013) adaptive control hypothesis, and has been used as evidence of a "perceptual wedge" in bilingual development (Petitto et al., 2012), as the metaphorical doorway to tuning to one’s native language is held open longer in bilingual development.

Not only is the influence of bilingual language experience on statistical learning abilities not yet clear, the relationship between an advantage in statistical learning abilities, executive function, and foreign language vocabulary learning is
even less well understood. Hence, the present dissertation aims to incorporate all three types of tasks in testing bilingual and monolingual participants, with the aim of further uncovering the interaction between these three types of cognitive skills.
CHAPTER 2: GOALS AND GENERAL DIRECTIONS

As reviewed in Chapter 1, the idea that there may be cognitive consequences and advantages for bilingualism is not a new one, but many aspects of such consequences are yet unknown, such as the roles that proficiency, language usage, age of acquisition, etc. play, as well as the underlying mechanism(s) driving these advantages. Additionally, the range of differential consequences that have been reported for bilinguals relative to monolinguals has been highly varied, in terms of the type of task or skill, the type of bilinguals who exhibit the effect, as well when in development particular advantages manifest (e.g., infancy, young adulthood, older adulthood). While it is beyond the scope of the present dissertation to address all of the questions that have been raised by the research reviewed, the present dissertation aims to answer some of these important questions. Specifically, the experiments reported here address the issue of whether bilinguals are advantaged learners generally, with an enhanced ability to acquire new information that exploits the same mechanisms that guide language processing.

2.1 Goals of the present research.

Across three experiments, the overarching research questions that the present dissertation aims to address are as follows: (a) How does knowledge of multiple languages affect various types of learning? (b) How do the specific properties of individual languages affect learning of a new language? And (c) What are the factors that predict language learning success? Addressing these questions not only has important theoretical implications for the scientific study of language
processing and language learning, but also important real-world consequences. Perhaps the most salient of these is the facilitation of language learning through a better understanding of the factors that can predict language learning success. Language learning is a crucial goal for individuals in a variety of populations, such as immigrants in the U.S. who struggle to learn English in order to be successful at school, in the workforce, and in service encounters such as banking and doctor’s appointments. For native-English speakers, learning a foreign language is often a criterion for employment, with knowledge of languages such as Japanese, Arabic, and Mandarin Chinese considered highly-valued skills. Knowing more about the mechanisms that underlie language learning will hopefully increase understanding about how to enhance success. In addition to the benefits these experiments may have for language learning, the results will help identify whether factors that predict language learning success also predict greater success in other types of learning. Such knowledge will allow for better testing and identification of learning outcomes, and with respect to language learning, it may allow assessment of outcomes before individuals have demonstrated behaviorally that they have learned any language at all.

The present dissertation reports three experiments comparing bilingual and monolingual performance across a variety of tasks in order to investigate the way in which past learning of one type affects new learning of the same or different type. The first goal of the present work was to uncover how knowledge of multiple languages affects multiple forms of learning. In all three experiments, bilingual and
functionally monolingual individuals were tested in various types of learning tasks. This comparison is of interest because recent studies have shown that bilingualism appears to have consequences not only for language processing but also for cognition more generally (e.g., Bialystok et al., 2009; Bialystok, 2005). Additionally, as reviewed in Chapter 1, previous research has demonstrated that bilinguals outperform monolinguals in learning foreign language vocabulary (e.g., Kaushanskaya & Marian, 2009a; 2009b). However, the relationship between these two skills is unknown, and it remains to be determined whether the advantage in foreign vocabulary learning is merely an extension of enhanced executive function.

In testing both linguistic and non-linguistic learning, the relationship between the bilingual advantage in executive function and in foreign vocabulary learning will be better understood. These results hold significance not only for understanding bilingualism and second language acquisition, but also for statistical learning and pattern recognition, processes that appear to identify individual differences that can predict learning success. Additionally, there is a suggestion that bilinguals are disadvantaged on tasks of lexical access relative to monolinguals, even when using the dominant language (e.g., Gollan et al., 2008). This work helps to disentangle the seemingly contradictory findings that bilinguals are slower than monolinguals to retrieve words in their dominant language, but also faster to retrieve words in a language just acquired.

The second goal of the experiments is to further investigate the role of bilingualism in new learning by manipulating the language used during study. As
reviewed in Chapter 1, previous research has demonstrated a bilingual advantage in foreign language vocabulary learning for bilinguals learning via L1 translations (Bogulski & Kroll, under review; Kaushanskaya & Marian, 2009a, 2009b; Kaushanskaya, 2012; Keshavarz & Astaneh, 2004; Papagno & Vallar, 1995; Sanz, 2000; Van Hell & Mahn, 1997). Additionally, research to date has only been able to demonstrate a bilingual advantage when bilinguals learn via L1 translations. However, to date, no studies investigating this particularly bilingual advantage have used ERPs to characterize the advantage. To this end, Experiment 1a tested monolinguals and bilinguals on a series of vocabulary learning tasks using both behavioral and ERP methods. Given previous evidence found in the literature on the bilingual advantage in foreign language vocabulary learning and on ERP vocabulary learning, several questions could be raised. For example, given a word/nonword judgment task, do bilingual learners exhibit an N400 to nonwords relative to real words at an earlier phase of learning? Do they exhibit a larger N400? Is it a combination of both? Or are bilingual and monolingual brains processing the words and nonwords similarly—at least, as indicated by the N400—and some additional mechanism is responsible for the advantage? Experiment 1b aimed to investigate these possibilities by administering a set of foreign language vocabulary learning and test tasks to a new group of bilingual participants, but they were asked to learn the unfamiliar vocabulary via L2 translations. These bilinguals were sampled from a bilingual population that had already been reported to exhibit an advantage in foreign language vocabulary learning via L1 translations (e.g. Bogulski & Kroll,
under review). The issue of potential limitations to this advantage is highly relevant for educational contexts, where non-native English speakers are often treated as a uniform group of language learners. Understanding the ways in which the languages involved in learning can affect the learning of a target language would be highly applicable finding. Some recent evidence has already demonstrated that an individual’s native language does influence language learning. Bogulski and Kroll found that the bilingual advantage in foreign vocabulary learning is restricted to bilinguals learning new words via L1 translations. This suggests that this bilingual advantage is not dependent on domain-general cognitive processes, but rather, is reliant on the route of translation via the L1 and the specific processes engaged with respect to each of the bilingual’s two languages.

The final overarching goal of the planned research is to gain insight into the factors that predict language learning success. Not only is this type of evidence highly pedagogically applicable, it can also add to an existing body of knowledge regarding how language processing proceeds at an early stage. Experiments 1a and 1b are the first to examine multilingualism as a predictor of learning as detected through ERP methodology. Not only does the introduction of ERP methods bridge the literatures of second language acquisition (and adapts a methodology utilized by McLaughlin et al., 2004, to examine group differences) and bilingual advantages in foreign vocabulary learning, it also allows for an examination of language learning between groups predicted to have differential learning success at earlier stages of learning than previously examined. The results may determine more precisely at
what stage of the learning process bilinguals begin to exhibit an advantage over monolinguals, which can further inform at what stage of learning individual differences can impact learning success, be it at an explicit, behavioral level or an implicit, neural one.

Experiment 2 tested bilinguals and monolinguals on two tasks: statistical learning and pattern recognition. The goal of this experiment was to determine whether success in one type of learning may or may not transfer to other types. For instance, it may be the case that the learning of foreign language vocabulary words may rely on an underlying process that is also exercised in statistical learning and/or pattern recognition. By comparing bilinguals and monolinguals, this experiment aimed to further elucidate the scope of the bilingual advantage.

2.2 General Predictions

Given the review of the relevant literature and the stated goals of the present dissertation, the following is a set of predictions for the stated experiments.

2.2.1 Experiment 1a Predictions. In Experiment 1a, English-Spanish bilinguals and English monolinguals learned unfamiliar Dutch vocabulary words and were tested on their ability to later recognize those words while ERPs were recorded. By testing bilinguals and monolinguals on a foreign language vocabulary learning task learning via L1 translations, the most obvious prediction made for Experiment 1a is that behaviorally, the bilinguals will prove to be better learners than will monolinguals. Experiment 1a presents multiple testing opportunities that are either immediately following the presentation of the study task or delayed by
approximately one week. However, previous foreign language vocabulary learning studies have found bilingual advantages for both immediate and delayed testing phases (e.g., Bogulski & Kroll, under review; Kaushanskaya & Marian, 2009a, 2009b; Van Hell & Mahn, 1997), and as such, bilinguals in Experiment 1a were expected to outperform monolinguals at all critical testing phases (Tests 1, 2, and 3).

More speculative are the predictions regarding the ERP data from Experiment 1a. McLaughlin et al. (2004) found increasing sensitivity to pseudo-French words for French learners in the form of greater amplitude of the N400 component in response to pseudo-French words relative to French words across testing sessions. However, Experiment 1a differs from the McLaughlin et al. study in a number of ways. First, the foreign language learning exposure in Experiment 1a was done in the laboratory, and was not based in the classroom. Second, due to the laboratory nature of the experiment, testing sessions were not conducted over the course of several months, but rather limited to two sessions separated by approximately one week. Finally, the group comparison conducted by McLaughlin et al. was between classroom learners and non-learners of French, whereas in Experiment 1a, the comparison was between monolinguals and bilinguals—or, in the terms utilized by Van Hell and Mahn (1997), novice and experienced learners.

Given the ways in which Experiment 1a and the McLaughlin et al. (2004) study differ, a number of possible outcomes might be considered. If it is the case that the bilingual advantage in foreign language vocabulary learning extends to early processing mechanisms indexed by the N400, then the bilingual group should
demonstrate enhanced sensitivity relative to the monolinguals in discriminating the Dutch-like nonwords relative to the real Dutch words. Because the bilingual advantage in foreign language vocabulary learning has been reported for both immediate and delayed testing phases, it is also then reasonable to predict that, if this advantage extends to a sensitivity realized at an electrophysiological level, this enhanced sensitivity for bilinguals relative to monolinguals may occur as early as Test 1. If this prediction is supported, it is also worth investigating whether this enhanced electrophysiological sensitivity is greatest at earlier testing phases relative to later ones. Previous research has not indicated that the magnitude of the bilingual advantage in foreign language vocabulary learning differs across immediate and delayed testing sessions, at least for behavioral data. However, as the McLaughlin et al. study has demonstrated, ERP data have the potential to reveal effects not present in the behavioral record. Hence, it is possible that sensitivity to the real Dutch words relative to the Dutch-like nonwords as indexed by the N400 will either remain constant across testing sessions (as predicted by the behavioral evidence), or vary across them (if the ERP evidence reveals differential effects not present in the behavioral record).

2.2.2 Experiment 1b Predictions. Like Experiment 1a, in Experiment 1b an additional group of English-Spanish bilinguals were asked to learn Dutch vocabulary and were later tested on their memory of these words while both ERP and behavioral measures were recorded. However, unlike Experiment 1a, the bilinguals tested in Experiment 1b learned Dutch vocabulary via L2 (Spanish) translations.
Based on the results reported by Bogulski and Kroll (under review), who found that only bilinguals who learned foreign language vocabulary via L1 translations exhibited an advantage relative to monolinguals, the behavioral data are predicted not to demonstrate an advantage relative for the bilinguals in Experiment 1b relative to the monolinguals from Experiment 1a, as the bilinguals in Experiment 1b learned the Dutch vocabulary via translations in their L2: Spanish. The critical test will be to compare the ERP data for the bilinguals learning via L2 vs. L1 translations. If ERP data can reveal evidence of learning above and beyond what behavioral data demonstrates, then it may be that bilinguals learning via L2 translations may exhibit enhanced sensitivity to the Dutch-like nonwords relative to the real Dutch words (as indexed by the N400) compared to the monolinguals from Experiment 1a. It may also be that similar neurophysiological effects can be observed for both bilingual learning via L1 translations and bilingual learning via L2 translations, but that bilinguals learning via L2 translations experience some other impediment to learning success not indexed by the N400. By contrast, it may be that bilinguals learning via L2 translations do not show enhanced sensitivity to the Dutch-like nonwords relative to the real Dutch words relative to monolinguals, as they may not be exercising the critical mechanism leading to advantaged foreign language vocabulary learning because they are learning via L2 translations. Hence, if learning via L1 translations is a necessary learning context for the bilingual advantage in foreign language vocabulary learning, then the bilinguals in Experiment 1b may
exhibit performance similar to that of the monolinguals in Experiment 1a, in both the behavioral and ERP record.

**2.2.3 Experiment 2 Predictions.** The goal of Experiment 2 was to compare bilingual and monolingual performance on a series of learning/pattern detection and executive function tasks (as well as several measures of individual differences) to determine if bilinguals have advantages relative to monolinguals for learning tasks generally. The two critical learning/pattern detection tasks administered in Experiment 2 were a statistical learning task adapted from Gebhart, Aslin, and Newport (2009) and a dot-motion discrimination task adapted from Green, Pouget, and Bavelier (2010). In the statistical learning task, subjects were passively exposed during a learning phase to a series of tri-syllabic words that conformed to one of two different artificial languages that differed in their syllables, but had the same underlying statistical properties. Critically, the two languages were presented one after the other without an explicit cue that there had been a change. After one exposure to both languages, subjects were tested on their ability to discriminate between words and nonwords in both languages. Subsequently, they were once more passively exposed to the two languages as in the first learning phase, but with the order of the languages reversed, and tested once more. In between the two statistical learning tasks, the participants also completed the Operation-Span task to measure working memory, the AX-CPT to assess inhibitory control, the Boston Naming test to assess English proficiency and lexical access, and the dot-motion discrimination task.
Both the statistical learning task and the dot-motion discrimination task
require implicit processing to extract patterns from a noisy input. In the statistical
learning task, a degree of learning is thought to have occurred when performance on
the two-alternative forced choice test is greater than 50%. Some research has
provided evidence that bilinguals may be advantaged on implicit learning tasks
(Bartolotti et al., 2011; Kovács & Mehler, 2009; Nation & McLaughlin, 1986),
although not all studies report such advantages (Grey, 2013; Yim & Rudoy, 2013). In
using multiple language streams, a la Gebhart et al. (2009), a comparison between
bilingual and monolingual performance can be made for initial language learning
(Language A, first learning phase) as well as language learning without an explicit
cue to signal a change in language (Language B, first learning phase). Additionally,
by re-administering the statistical learning task in the reverse order (i.e., Language
B followed by Language A), bilingual and monolingual performance can be
compared for retention after intervening exposure to a different language
(Language A, second learning phase), as well as re-learning of a language previously
encountered without a boundary cue signaling the change in language (Language B,
second learning phase). Hence, while bilinguals may not outperform monolinguals
on all statistical learning tests, they may be better able to detect a change in
language input without an explicit cue. This type of advantage would emerge as a
diminished primacy effect, or perhaps an ability to learn both Language A and
Language B in the first learning exposure without an explicit cue signaling the
change from one to the other, whereas the monolinguals may be able to learn only
Language A (as reported by Gebhart et al.). Along these lines, if the bilinguals are better able to attend to the implicit cues signaling a language change in the first exposure of the two languages, they may experience enhanced learning during the second exposure of the two languages for both languages, relative to the first exposure, as they would be better able to, again, attend to the implicit cues denoting a language change, and would hence outperform the monolinguals who were less adept at learning due to the lack of an explicit cue.

The dot-motion discrimination task has previously been reported as a task on which action video game players outperform non-players. This task involves exposing participants to an array of moving dots, most of which are moving in a random direction. The remaining dots (a variable percentage) are moving either left or right. Subjects simply, but quickly, have to decide whether the coherent direction of motion of the dots is to the left or right. Similar to the statistical learning task, skilled performance on this task is thought to index enhanced probabilistic inference. However, unlike the statistical learning task, the attentional demands of the dot-motion discrimination task are much greater, as responses are required within a 2s time window for each trial, and detecting lower levels of coherent motion among the dots is especially difficult even without time pressure. Additionally, the statistical learning task utilized in Experiment 2 was auditorily presented, while the dot-motion discrimination task was entirely visual. Hence, if bilinguals have an enhanced mechanism that is exercised on tasks that require skilled statistical learning (which may simply be another term for probabilistic
inference, albeit from a different research literature), then the bilinguals should outperform the monolinguals on both the statistical learning and dot-motion discrimination task. This result would support the idea that bilinguals are advantaged at multiple types of new learning, particularly in implicit processing that facilitates statistical learning of transitional probabilities as well as in an attentionally-demanding probabilistic inference task.

If, on the other hand, bilinguals do not demonstrate advantages on both tasks, then implicit processing would not be a domain in which bilinguals outperform monolinguals. Rather, if bilinguals exhibit an advantage relative to monolinguals on only the dot-motion discrimination task, this would further support claims of a bilingual advantage in executive function, and particularly in the attentional component, as the attentional demands of this task are quite high. If this were the case, then this would suggest that the bilingual advantage in foreign language vocabulary learning is not simply another manifestation of the bilingual advantage in implicit learning/processing. If an advantage existed for bilinguals on the statistical learning task to the exclusion of the dot-motion discrimination task, this would add support to the idea that bilingual advantages in implicit learning may be particular to linguistic domains, or perhaps are exclusive to contexts that do not place great demands on attentional networks. This would imply that the bilingual advantage in foreign language vocabulary learning is not rooted in an underlying advantage in implicit processing, but is more specific, and depends on the contextual environment of learning.
Finally, performance on the AX-CPT may be related to language experience (i.e., being bilingual or monolingual), as well as to performance on the statistical learning and/or dot-motion discrimination task. The AX-CPT used in Experiment 2 (with distractors) is a task that exposes participants to rapidly presented streams of 5 letters, the first of which is a cue, the last of which is a target, and the middle three of which are distractors. Their task is to press a NO key for each distractor and cue, and also for the target under certain combinations of cue and target. Thus, it involves preparatory responses as well as the inhibition of those responses. As described by Morales et al. (in press), proactive inhibitory control is thought to be more or less synonymous with monitoring or goal maintenance, whereas reactive control can be described as response suppression. The AX-CPT is thought to require both types of inhibitory control, and may further disentangle which inhibitory processes bilinguals exercise differentially relative to monolinguals. The results of the Morales et al. experiment suggest that the bilinguals in Experiment 2 should outperform monolinguals for trials that emphasize the role of proactive inhibitory control in particular. However, Morales et al. tested early bilinguals (i.e., bilinguals who had learned a second language before the age of 6), so Experiment 2 asks whether a group of late bilinguals performing the same task exhibit similar advantages in proactive inhibition as the early bilinguals reported by Morales et al. If these results were replicated in Experiment 2, it may also follow that better proactive inhibitory processing may facilitate performance on either/both the
statistical learning or dot-motion discrimination task, as both tasks may rely on skilled inhibitory processes for success.
CHAPTER 3: EXPERIMENT 1A – USING ERPs TO INVESTIGATE THE BILINGUAL ADVANTAGE IN FOREIGN LANGUAGE VOCABULARY LEARNING

3.1 Experiment 1a

3.1.2 Method.

3.1.2.1 Participants. Twenty-five functionally monolingual English speakers and 23 native-English speakers who had learned Spanish as a second language were recruited for Experiment 1a. Data from two monolingual participants were discarded due to experimenter error, and data from three monolingual participants and four bilingual participants were discarded due to a lack of a sufficient number of trials after artifact rejection per condition for inclusion in the ERP analysis. Finally, data from one bilingual participant were discarded due to Spanish as a home language. The results presented for Experiment 1a reflect the remaining data from 20 monolingual and 18 bilingual participants. The English-Spanish bilinguals recruited for this study were predominantly late learners of Spanish (L2 onset: $M = 12.5$; range: 4 to 18, with only 2 of the 18 participants first exposed to Spanish earlier than age 12).

All participants were recruited from either the campus or the outlying areas surrounding the Pennsylvania State University and were compensated $10 per hour for their participation. As Experiment 1a consisted of two sessions approximately
one week apart (range: 5 to 9 days, \( M = 6.9 \) days), participants were compensated for their time at the end of each session.

### 3.1.2.2 Materials and Design

The study and test tasks consisted of 90 Dutch words that were selected for their lack of orthographic overlap with English (i.e., they were neither cognates nor false cognates). The properties of these Dutch words and their English translations can be found in Appendix. The words chosen for inclusion in the study varied across a number of attributes, including word class (nouns, verbs, adjectives, and adverbs were represented), length, frequency, etc., in order to reflect a diverse set of items that a real-life learner of Dutch might encounter. A set of 90 Dutch-like nonwords were created for use in the Dutch lexical decision task using the Pseudo computer program (Van Heuven, 2003), using the original Dutch items to generate nonwords matched in length, bigram and trigram frequencies, and consonant-vowel structure. The nonwords were checked against both English and Dutch dictionaries to confirm their nonword status. The nonwords were repeated throughout the four test task presentations in order to minimize novelty effects in the ERP record that could have obscured effects of learning (though this created a potential problem of the nonwords becoming more "word-like" with repeated exposure, especially within an experimental session; see the discussion in section 3.1.4).

Both the English and Spanish picture naming tasks consisted of the same 66 high frequency and 66 low frequency pictures utilized by Gollan, Montoya, Cera, and Sandoval (2008).
3.1.2.3 Procedure. The experiment consisted of two sessions separated by approximately one week. A schematic of Experiment 1a can be found in Table 3.1.

Table 3.1: Experiment 1a Procedure Overview. * denotes that ERP data were collected.

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Pre-Test</td>
<td>*Test 2</td>
</tr>
<tr>
<td>*Study Task 1</td>
<td>*Study Task 3</td>
</tr>
<tr>
<td>Executive Function Tasks</td>
<td>English Picture Naming</td>
</tr>
<tr>
<td>*Study Task 2</td>
<td>*Test 3</td>
</tr>
<tr>
<td>*Test 1</td>
<td>Spanish Picture Naming (Bilinguals Only)</td>
</tr>
</tbody>
</table>

Experiment 1a consisted of multiple study and test tasks to teach and assess learning of the 90 Dutch words and their English translations, as well as several measures of executive function and language proficiency. All participants were tested in a quiet chamber, using E-Prime software on a PC computer for display and behavioral data acquisition. All tasks for which EEG data was collected (i.e., test and study tasks) were presented with white text on a black background to minimize eye strain. The other tasks were presented with black text on a white background. When possible, the lights were dimmed during tasks for which EEG data was collected and brightened during behavior-only tasks. Before each experimental session began, participants watched the live recording of their EEG signal, and were asked to blink, clench their jaws, move their eyes side to side, and move their bodies; all of these behaviors generate noise artifacts in the EEG signal. In order to minimize data loss from such noise, participants were also instructed to find a comfortable position before the tasks began, so as to restrict their body movements as much as possible.
during the EEG recording. Additionally, to restrict EEG data loss due to blink artifacts, participants were strongly encouraged to restrict their blinking to when the fixation cross was presented, and to try to avoid blinking when items were presented.

3.1.2.3.1 Recognition Test Task. The test task was similar to a lexical decision task, consisting of all 90 Dutch words the participants were being taught, as well as the 90 matched Dutch-like nonwords. Because these participants were by no means proficient Dutch speakers, the label "lexical decision" is not meaningful, as some of the critical assumptions underlying a lexical decision task (e.g., a decision can sometimes be made without accessing the lexicon) do not hold for these participants. Henceforth, this task will be referred to simply as the recognition test task or simply the test task.

The presentation of the Dutch words and the Dutch-like nonwords in the test task was pseudorandomized, such that no more than three items of a given type (i.e., word or nonword) appeared consecutively. A trial consisted of the following sequence: a fixation cross (displayed for 2 s), a blank screen (displayed for 350 ms), and the critical item (either a Dutch word or a Dutch-like nonword; displayed until the participant responded, up to 5 s). The participants were instructed to respond with a key press as quickly and accurately as possible whether the item presented on the screen was a word or a nonword. Each presentation of the test task had a break after half of the critical items had been presented, during which time the experimenter brightened the lights (if dimmed) and engaged the participant in
conversation for a short period of time, in order to combat drowsiness and alpha wave artifacts.

The first presentation of the test task occurred before the participant had been presented with the study task, and served as a pre-test. This pre-test served as a baseline measure both for behavioral performance and ERP comparisons within and across participants.

3.1.2.3.2 Study Task. The study task consisted of the 90 tested Dutch words presented with their correct English translations. A given trial of the study task proceeded as follows: a double fixation cross (displayed for 350 ms; one + closely above the other, in an attempt to minimize saccades), a blank screen (displayed for 300 ms), a Dutch word replacing the upper fixation cross (displayed for 2 s), the Dutch word remaining on screen while the lower fixation cross disappears (displayed for 300 ms), the English translation of the previous Dutch word replacing the lower fixation cross (displayed until the participant named the English translation aloud, up to 5 s), followed by a blank screen (displayed for 300 ms). Participants were instructed to take as long as they needed within the 5 second time window to name the English translation, as this would cue the test task to advance. Additionally, participants were told that they would never be required to pronounce the Dutch words, only recognize them.

3.1.2.3.3 English and Spanish Picture-Naming Tasks. Both the English and Spanish picture naming tasks were administered using the procedure outlined by Gollan et al. (2008). A given trial in either the English or the Spanish picture naming
task proceeded as follows: a fixation cross was presented until the participant pressed a key to initiate the trial (display under participant control), followed by a blank screen (displayed for 350 ms), followed by a black and white line drawing (displayed until a voice-key was activated by the participant naming the picture into the microphone, up to 3 s). The participants were instructed to name pictures as quickly as possible without making mistakes, and to avoid making additional noises if possible, which could inadvertently cue the voice key. The pictures were pseudorandomly mixed, such that no more than three pictures of a given frequency category (i.e., high or low frequency) appeared consecutively.

3.1.2.3.4 Flanker Task. A Flanker task (as utilized by Emmorey, Luk, Pyers, & Bialystok, 2008) was administered in order to assess multiple aspects of executive function, and the Operation-Span task (Turner & Engle, 1989) was administered in order to assess working memory capacity. These two tasks were presented in between the first and second presentation of the study task in the first session of Experiment 1a. As in Emmorey et al. (2008), participants were presented with displays of chevrons, diamonds, and Xs (see Figure 3.1).
Control blocks consisted of control trials only, which were the presentation of a single red chevron pointing either left or right. The participant was instructed to respond to the chevron’s direction as quickly and accurately as possible. In the congruent/incongruent blocks, participants were again asked to respond as quickly and accurately as possible to the direction of the red chevron, which was now flanked by either black chevrons pointing in the same direction as the target (congruent trials) or black chevrons pointing in the opposite direction as the target.
(incongruent trials). Go/no-go blocks consisted of trials with either black diamonds of black Xs flanking a red chevron. During these blocks, participants were instructed to respond as quickly and accurately to the directly of the red chevron when the red chevron was flanked by black diamonds (go trials), and not to respond at all when the red chevron was flanked by black Xs (no-go trials). For congruent, incongruent, go, and no-go trials, the red chevron could appear in either the second, third, or fourth position in the five-item sequence, pointing either left or right. In addition to these blocks, there were two mixed blocks consisting of congruent, incongruent, go, and no-go trials intermixed in a random but fixed presentation. The Flanker task began and ended with control trial blocks (consisting of 12 control trials), and two mixed blocks in the middle of the task (consisting of 36 trials each, 9 trials per condition). The congruent/incongruent and go/no-go blocks were presented between the control block and the mixed blocks, and again in reverse order between the mixed blocks and the final control block. The order of the congruent/incongruent and go/no-go blocks was counterbalanced between participants.

3.1.2.3.5 Operation-Span Task. The Operation-Span task consisted of 60 trials, each of which paired an arithmetic equation (half of which were correct, half of which were incorrect) followed by a word in English. The participants were instructed to respond whether the equation presented was correct or incorrect as quickly and accurately as possible (with a time limit of 2.5 seconds) with a key press, and then to commit to memory the word following the equation. Following
each block of equation/word pairs, participants were cued to type all of the words they could remember from the block. The equation/word pairs were divided into blocks of two, three, four, five, or six, with three blocks of two pairs presented first, followed by three blocks of three pairs, then three blocks of four pairs, etc. All participants were instructed to do their best on both solving the equations quickly and accurately as well as memorizing the words.

3.1.2.4 EEG acquisition and offline processing. Continuous EEG signal was recorded from each participants using a 32-channel sintered Ag/AgCL electrode array in an elastic cap according to the 10-20 system (QuikCap, Neuroscan Inc.; see Figure 3.2).
Figure 3.2. 32-channel QuikCap electrode montage used for ERP data collection.

This signal was amplified using Neuroscan Synamps with a band pass filter of 0.05 to 100 Hz, with a sampling rate of 500 Hz. The recordings were referenced online to the right mastoid site, and re-referenced to the left mastoid site offline during post-processing of the EEG data. Electrical impedances at all electrode sites were brought
at or below 5 kΩ before recording on the first task in a given session, were checked at the end of each task and, if necessary, brought back down to that level at the start of the next task recording continuous EEG. Six loose electrodes were affixed to the following sites for each participant: left mastoid (for offline referencing), right mastoid (for online referencing), the two outer canthi of both eyes (to track horizontal eye movements), and the upper and lower orbital ridges of the left eye (to track vertical eye movements). Recordings from the eye channels were used in post-processing of the EEG data to identify trials in which participants blinked or exhibited saccadic eye movement; these trials were discarded from analysis.

The post-processing of the continuous EEG recordings proceeded as follows. First the data were imported into Matlab for processing and analysis using the EEGLAB toolbox (Delorme & Makeig, 2004) and the ERPLAB plugin (Lopez-Calderon & Luck, 2010). Then, all channels were re-referenced to the left mastoid (excepting the eye channels and the mastoids themselves). Next, the continuous data was re-filtered using a second-order infinite impulse response (IIR) Butterworth filter to filter out all signal below a threshold of 19.4 Hz. Following the application of the offline low-pass filter, epoched data were extracted from the continuous EEG signal according to condition (i.e., word or nonword trials), using time windows of -200 to 800 ms surrounding the presentation of a stimulus. A baseline correction was then applied to each extracted epoch according to a pre-stimulus baseline of an averaged 200 ms before the presentation of an item to the moment of presentation. Next, two artifact detection methods were applied in parallel: a simple voltage threshold of
-100μV to 100μV inclusion criterion across all electrode sites, and a moving window peak-to-peak threshold of 50μV with a window width of 200 ms for the frontal electrodes only (C4, T8, FP1, CZ, FC4, FT8, C3, FZ, F4, F8, T7, FT7, FC3, F3, and FP2). Participants with fewer than 13 artifact-free trials (< 30%) in any condition at any presentation of the test task were excluded for all analyses. Analyses on the ERP test task data were repeated-measure ANOVAs, using condition (two levels: nonword, word) and test task presentation order (four levels), carried out separately for monolinguals and bilinguals. Midline analyses also included electrode as an additional within-subjects factor with three levels: FZ, CZ, and PZ. Medial-lateral analyses included both electrode (five levels) and hemisphere (two levels: left, right) as additional within-subject factors, and included FP1, F3, C3, P3, and O1 as left hemisphere electrodes, and FP2, F4, C4, P4, and O2 as right hemisphere electrodes. Lateral-lateral analyses also included both electrode (three levels) and hemisphere (two levels: left, right) as additional within-subject factors, and included F7, T7, and P7 as left hemisphere electrodes, and F8, T8, and P8 as right hemisphere electrodes. The Huynh-Feldt correction was used for all effects with $df > 1$ in the numerator to correct for violations of sphericity, but uncorrected degrees of freedom are reported.

Due to the nature of the research question regarding the possibility of brain response differentiation between words and nonwords in the absence of behavioral differentiation, both correct and incorrect trials were included in the ERP analyses. The N400 component was selected as the component of interest for analysis.
purposes, due to previous experimental work identifying this component as one that may index lexical retrieval, even in early learners (e.g., McLaughlin et al., 2004; Midgley, Holcomb, & Grainger, 2009). Upon visual inspection of the data, the N400 time window was identified as occurring between 300 and 600 milliseconds following the presentation of a stimulus. Mean amplitude in the 300 to 600 millisecond time window was the dependent measure.

3.1.3 Results: Behavioral Data. The goal of Experiment 1a was twofold: first, to replicate the finding of a bilingual advantage in foreign language vocabulary learning in late bilinguals as reported by Bogulski and Kroll (under review); and second, to examine whether such an advantage can be observed at a neurophysiological level as well, as operationalized as a larger N400 at centro-parietal sites to Dutch-like nonwords relative to studied Dutch words and/or earlier evidence of an N400 at such sites to nonwords relative to words. As such, the results of Experiment 1a will be presented and discussed for the behavioral data first, and then re-examined in the context of the ERP data.

Participant descriptive statistics, including Operation-Span score, mean naming latencies and accuracy on the English version of the Boston Naming test, can be found in Table 3.2 below.
Table 3.2. Mean age, age of L2 onset, L1 self-ratings, L2 self-ratings, Operation-Span score, and mean naming latencies and accuracy on the Boston Naming test for participants in Experiment 1a (standard deviations in parentheses, significant differences between groups are denoted with *; marginal significance is denoted with †).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age</th>
<th>Days Between Session 1 &amp; 2</th>
<th>Age of L2 Onset†</th>
<th>L1 Self-Rating</th>
<th>L2 Self-Rating†</th>
<th>O-Span Score†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolinguals</td>
<td>20</td>
<td>22.1</td>
<td>7.1</td>
<td>14.3</td>
<td>9.8</td>
<td>2.9</td>
<td>44.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.7)</td>
<td>(0.4)</td>
<td>(2.1)</td>
<td>(0.3)</td>
<td>(2.0)</td>
<td>(8.8)</td>
</tr>
<tr>
<td>Bilinguals</td>
<td>18</td>
<td>21.8</td>
<td>6.8</td>
<td>12.5</td>
<td>9.6</td>
<td>7.1</td>
<td>49.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.7)</td>
<td>(0.7)</td>
<td>(3.3)</td>
<td>(0.4)</td>
<td>(1.6)</td>
<td>(6.4)</td>
</tr>
</tbody>
</table>

No differences existed between the two groups for age ($F < 1$), the number of days between Sessions 1 and 2 ($F(1, 36) = 1.40, p = .25$), or L1 self-ratings ($F(1, 36) = 1.92, p = .17$). The difference in age of L2 onset was marginally significant ($F(1, 36) = 3.74, p = .06$), as the bilinguals overall had their first exposure to a second language earlier than did the monolinguals. Additionally, bilinguals rated themselves as much more proficient in an L2 than did the monolinguals ($F(1, 36) = 51.57, p < .001$).

Finally, there was a marginally significant difference in working memory between the groups, as the bilinguals outperformed the monolinguals on the Operation Span task ($F(1, 36) = 4.09, p = .051$).

**3.1.3.1 Flanker task.** The accuracy rates for both bilinguals and monolinguals on all trial types in the flanker task was very high (all group averages
per condition were at least 90%). Due to these high levels of accuracy, the accuracy data will not be analyzed any further.

The reaction time data were analyzed using five separate ANOVAs. First, the Flanker Effect was calculated as the speed of response to incongruent trials relative to congruent trials, in an experimental block containing only those two types of trials. A one-way ANOVA revealed no differences between bilinguals and monolinguals in the Flanker Effect ($F(1, 36) = 1.00, p = .32$). Next, speed of performance for each trial type (i.e., incongruent, congruent, go, and control) in the blocked condition was compared using a 2 (group: monolinguals, bilinguals) by 4 (trial type: incongruent, congruent, go, and control) mixed-factor ANOVA. The results showed no main effect of group, nor a group by trial type interaction (both $F$s < 1), but did reveal a main effect of trial type ($F(3, 108) = 152.11, p < .001, \eta^2_p = .81$). Planned Bonferroni-corrected pairwise comparisons revealed that control trials were faster than all other trial types (all $p$s < .001), followed by congruent trials, which were faster than both go and incongruent trials (both $p$ values < .001). No difference between go and incongruent trials was found.

Next, a 2 (group: monolinguals, bilinguals) by 3 (trial type: incongruent, congruent, go) mixed-factor ANOVA was conducted on the reaction time data for those trials that were in the mixed block, which contained incongruent, congruent, go, and no-go trials. The results again demonstrated no main effect of group ($F(1, 36) = 1.19, p = .28, \eta^2_p = .03$), nor a group by trial type interaction ($F < 1$). A main effect of trial type did emerge ($F(2, 72) = 58.24, p < .001, \eta^2_p = .62$), and as planned
Bonferroni-corrected pairwise comparisons revealed, the congruent trials were significantly faster than go trials, which were significantly faster than incongruent trials (all $p$ values < .01).

A fourth ANOVA was conducted on the mixing cost data, which were difference scores calculated by subtracting reaction times for incongruent, congruent, and go trials in the blocked condition from the reaction times for those same trials in a mixed block. Hence, a 2 (group: monolinguals, bilinguals) by 3 (trial type: incongruent, congruent, go) mixed-factor ANOVA revealed no main effect of group ($F(1, 36) = 1.88, p = .18, \eta_p^2 = .05$), but did reveal a marginally significant group by trial type interaction ($F(2, 72) = 2.89, p = .07, \eta_p^2 = .07$), as the two groups experienced similar costs for both congruent (monolinguals, 21ms; bilinguals, 22 ms) and incongruent (monolinguals, 47ms; bilinguals, 41ms) trials, but while monolinguals experienced a mixing cost for go trials as well (18ms), bilinguals did not (-14ms). Again, a main effect of trial type was observed ($F(2, 72) = 16.53, p < .001, \eta_p^2 = .32$), and planned Bonferroni-corrected pairwise comparisons revealed that overall, the greatest cost of mixing trial types was for the incongruent trials, followed by the congruent trials, followed by the go trials ($p$ values for all comparisons < .05).

Finally, an additional 2 (group: monolinguals, bilinguals) by 3 (trial type: incongruent, congruent, go) mixed-factor ANOVA was conducted on the relative costs for the incongruent, congruent, and go trials. Relative costs are calculated by comparing the speed of response for each of the critical trials to the control trials to
assess how much of an additional processing cost each type of trial imposes. This analysis revealed no main effect of group nor a group by trial type interaction (both $F$s < 1), though a main effect of trial type was again significant ($F(2, 72) = 27.25, p < .001, \eta_p^2 = .43$). Planned Bonferroni-corrected pairwise comparisons revealed that costs to both incongruent and go trials were greater than the cost to congruent trials (both $p$ values < .001), but there was no difference in costs to incongruent and go trials ($p = 1.0$).

The only group difference observed in the flanker task was an advantage for the bilingual group in the cost to mixing types of trials to go trials. Despite that the bilinguals tested by Luk (2008) were balanced bilinguals who had learned two languages from a very young age, relative to the late-learners tested in Experiment 1a, both studies found an advantage for the bilinguals in the mixing costs for go trials only, though statistical significance in Experiment 1a was marginal. However, Luk also found a bilingual advantage for the relative costs to incongruent trials (i.e., smaller values for incongruent trials relative to control trials), whereas the bilinguals in Experiment 1a did not exhibit this advantage. Hence, it may be the case that utilizing two languages more generally (independent of age of acquisition and perhaps even proficiency) may confer some benefit to cognitive flexibility, as it was in the mixing costs that the bilinguals outperformed the monolinguals. As this benefit was exclusive for the go trials, this benefit may also be related to monitoring in a cognitively demanding environment.
3.1.3.2 Word naming at study: Behavioral data. Overall, the English monolinguals correctly produced the target English translation during the study task 97.0%, 96.8%, and 98.2% of the time, for each of the three presentations of the study task, respectively. The English-Spanish bilinguals correctly produced the English translation during the study task 97.4%, 96.6%, and 96.4% of the time, for each of the three presentations of the study task respectively. The remaining trials were either trials named incorrectly (misread/mispronounced, or said the Dutch word aloud), or not named at all. This high level of accuracy is hardly surprising, as the task required participants only to read the English translation aloud when it was presented on the computer screen. Due to the high level of accuracy for all participants across all presentations of the study task, the accuracy data will not be analyzed further.

For the naming latency analyses, only trials identified as named correctly with accurate voice-key triggering were included. These naming latencies for both the monolingual and bilingual groups for all three presentations of the study task are presented in Figure 3.3.
A 2 (group: monolinguals, bilinguals) by 3 (study task presentations) mixed-factor ANOVA revealed no main effect of group ($F(1, 36) = 1.90, p = .18$) or of presentation order ($F(2, 72) = 0.69, p = .51$). It did, however, reveal a marginally significant group by presentation order interaction ($F(2, 72) = 2.75, p = .07$), as bilinguals were slower to name the English translations aloud during the study task than were monolinguals, but this difference was larger at the latter two study presentations than the first.

Similar to results reported by Bogulski and Kroll (under review), the bilingual learners in Experiment 1a were slower in their English translation naming latencies during the study task, although this was only the case for the latter two presentations of the study task. The two studies were slightly different in their procedure, however, as the bilinguals in Experiment 1a had previously been
exposed to these words through the Pre-Test, which Bogulski and Kroll did not administer. Bogulski and Kroll argued that the increased naming latencies for bilinguals relative to monolinguals were due to increased interference from the English translations themselves in an L2 (or L3) context. They further argued that it was this interference and requisite inhibition that ultimately led to better learning in the English-Spanish bilinguals, relative to the monolinguals. Hence, as the bilinguals in Experiment 1a generally upheld this general pattern, it is reasonable to predict that they would also demonstrate an advantage in foreign language vocabulary learning relative to monolinguals.

3.1.3.3 Recognition test task: Behavioral data. As mentioned in the procedure (and as can be seen in the procedure represented in Table 3.1), the recognition test task (presented as a Dutch lexical decision task) was presented four times: twice in the first session (once as a pre-test, and once at the end of the session), and twice in the second session (once at the beginning, and once toward the end of the session). In order to account for potential response biases in the data, accuracy on the test task was converted to d’ scores. The data from these four test tasks for both the English monolinguals and the English-Spanish bilinguals are presented in Figures 3.4 (accuracy) and 3.5a/3.5b (reaction times).
A 2 (group: monolinguals, bilinguals) by 4 (test task presentations) mixed-factor ANOVA performed on the $d'$ data revealed a main effect of group ($F(1, 36) = 4.68, p < .05$), as the bilinguals were more accurate than the monolinguals on the test task overall. Additionally, a main effect of task presentation order emerged ($F(3, 108) = 4.40, p < .05$). Planned Bonferroni-corrected pairwise comparisons revealed that participants were significantly more accurate at all critical tests than at the pre-test (pre-test to test 1: $t1(36) = 11.12, p < .001$; pre-test to test 2: $t1(36) = 9.09, p < .001$; pre-test to test 3: $t1(36) = 10.81, p < .001$). Additional planned Bonferroni-corrected pairwise comparison further revealed that participant performance at test 1 was more accurate than at test 2 ($t1(36) = 8.45, p < .001$), but not at test 3 ($t1(36) = 1.15, p = 1.0$). Finally, performance at the final test was more accurate than performance at test 2 ($t1(36) = 7.19, p < .001$).
Hence, as Bogulski and Kroll (under review) found, and as expected given the longer naming latencies during the study task, the bilinguals in Experiment 1a outperformed the monolinguals at every iteration of the test task.

Figure 3.5a: Test task reaction times on nonword trials for Experiment 1a.
A 2 (group: monolinguals, bilinguals) by 4 (test task presentations) by 2 (condition: nonword, word) mixed-factor ANOVA performed on the reaction time data for the test task in Experiment 1a revealed no main effect of group ($F(1, 35) = 1.49, p = .23$), but there was a significant group by task presentation interaction ($F(3, 105) = 5.98, p < .01$), a marginally significant group by condition interaction ($F(1, 35) = 3.13, p = .09$, and a significant three-way interaction between group, task presentation, and condition ($F(3, 105) = 3.21, p < .05$). Reminiscent of the accuracy analysis, bilinguals were slower than monolinguals overall to perform the task at all presentations except for the pre-test. Additionally, monolinguals responded on word and nonword trials at almost identical speeds (1013 ms for nonwords and 1012 ms for words), whereas bilinguals were slower to respond to nonword trials (1144 ms for nonwords and 1074 ms for words). The ANOVA also revealed a marginally significant main effect of condition ($F(1, 35) = 3.27, p = .08$), as nonwords were slightly slower than word trials overall. This main effect is qualified by an additional condition by presentation order interaction ($F(3, 105) = 11.98, p < .001$), as words were slower than nonwords at the pre-test and at test 2, but nonwords were slower than words at test 1 and test 3. Finally, a main effect of task presentation order emerged ($F(3, 105) = 18.76, p < .001$), and planned Bonferroni-corrected pairwise comparisons revealed that pre-test reaction times were slower than responses at test 1 ($t1(35) = 2.96, p < .05$), test 2 ($t1(35) = 3.24, p < .05$), and test 3 ($t1(35) = 5.53, p < .001$). Additionally, performance at test 1 was faster than at test 3 ($t1(35) = 5.45$, $p < .001$).
$p < .001$), but similar in speed to performance at test 2 ($t(35) < 1, p = 1.0$). Finally, performance at test 2 was slower than at test 3 ($t(35) = 6.14, p < .001$).

It is unsurprising that performance at test 1 was faster than at the pre-test, given that by test 1, participants had previously been given two opportunities to study the Dutch words. Additionally, given that approximately one week elapsed between test 2 and test 3, it is also unsurprising that participants were slower and less accurate at test 3 relative to their performance at test 2. It also seems possible that the two groups were engaging in differential processing during the test task, as the bilinguals were much more consistent in their response times to the nonword trials than were the monolinguals, though the two groups were more similar in their speed of responses to the word trials. Additionally, except for the Pre-Test, the bilinguals were slower to perform the task overall, further supporting the idea that the bilinguals and monolinguals were performing the task differently.

Given that the bilingual advantage in foreign language vocabulary learning was replicated in the behavioral data in Experiment 1a for all critical tests, it is worthwhile to examine the ERP data collected at these tests to see whether the brains of the monolingual and bilingual learners responded differently to the Dutch words and nonwords. While the bilinguals exhibited an advantage relative to the monolinguals in their test task performance (as indicated by greater accuracy), it remains possible that both groups might show similar electrophysiological sensitivity to the nonwords relative to the words. In order to investigate possible
group differences, separate analyses were conducted on the ERP test task data for each group, and results were then compared between the two.

3.1.4 Results: ERP Data.

3.1.4.1 Test task ERP Results: Monolinguals. After being subjected to offline post-processing as described in section 3.1.2.4, and after the exclusion of all data from participants with fewer than 13 trials remaining in any condition at any test task presentation after artifact rejection, 74.02% of trials remained in the monolingual data for the test task analyses. A 2 (condition: nonword, word) by 4 (test task presentations) within-subjects ANOVA conducted on the percentage of trials retained revealed no differences in condition ($F(1, 19) < 1$), and no interaction between condition and task presentation ($F(3, 57) < 1$). While there was a marginally significant effect of time ($F(3, 57) = 2.67, p = .08$), this was driven by an increase in the number of rejected trials at the end of a session relative to the beginning, likely due to fatigue.

The average waveforms plotted for representative electrodes F3, FZ, F4, C3, CZ, C4, P3, PZ, and P4 for all four test task presentations (Pre-Test, Test 1, Test 2, and Test 3) can be found in Figures 3.6a, 3.6b, 3.6c, and 3.6d, respectively.
Figure 3.6a. Grand average ERP waveforms for monolinguals at the Pre-Test for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 3.6b. Grand average ERP waveforms for monolinguals at Test 1 for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during Test 1. μV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 3.6c. Grand average ERP waveforms for monolinguals at Test 2 for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during Test 2. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 3.6d. Grand average ERP waveforms for monolinguals at Test 3 for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during Test 3. μV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Monolinguals
Experiment 1a
English Translations

Pre-Test

Test 1

Test 2

Test 3

-150 0 300 700
ms

-5 0 5
μV

Nonwords
Words
Figure 3.7. Grand average ERP waveforms for monolinguals from Experiment 1a at representative electrode site CZ for all test task presentations. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed line found at 0 ms represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.

As previously mentioned, the ERP data from the monolingual group was analyzed using three separate repeated-measures ANOVAs for midline electrodes, medial-lateral electrodes, and lateral-lateral electrodes. The results of these analyses are presented in Table 3.3.
Table 3.3. Monolingual ERP ANOVA results for Experiment 1a.

<table>
<thead>
<tr>
<th></th>
<th>F-value</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 19)</td>
<td>3.68*</td>
<td>0.16</td>
</tr>
<tr>
<td>Test Task Presentation (3, 57)</td>
<td>7.73***</td>
<td>2.89</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 57)</td>
<td>&lt; 1</td>
<td>0.05</td>
</tr>
<tr>
<td>Electrode (3, 57)</td>
<td>20.23***</td>
<td>0.52</td>
</tr>
<tr>
<td>Condition x Electrode (3, 57)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (9, 171)</td>
<td>5.83***</td>
<td>0.24</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (9, 171)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Medial-Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 19)</td>
<td>2.4</td>
<td>0.11</td>
</tr>
<tr>
<td>Test Task Presentation (3, 57)</td>
<td>4.39**</td>
<td>0.19</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 57)</td>
<td>&lt; 1</td>
<td>0.05</td>
</tr>
<tr>
<td>Electrode (6, 114)</td>
<td>10.44***</td>
<td>0.36</td>
</tr>
<tr>
<td>Condition x Electrode (6, 114)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (18, 342)</td>
<td>1.76</td>
<td>0.09</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (18, 342)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td>Hemisphere (1, 19)</td>
<td>27.46***</td>
<td>0.59</td>
</tr>
<tr>
<td>Condition x Hemisphere (1, 19)</td>
<td>1.69</td>
<td>0.08</td>
</tr>
<tr>
<td>Test Task Presentation x Hemisphere (3, 57)</td>
<td>5.02**</td>
<td>0.21</td>
</tr>
<tr>
<td>Electrode x Hemisphere (6, 114)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Hemisphere (3, 57)</td>
<td>&lt; 1</td>
<td>0.01</td>
</tr>
<tr>
<td>Condition x Electrode x Hemisphere (6, 114)</td>
<td>&lt; 1</td>
<td>0.02</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode x Hemisphere (18, 342)</td>
<td>1.24</td>
<td>0.06</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode x Hemisphere (18, 342)</td>
<td>1.07</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Lateral-Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 19)</td>
<td>1.69</td>
<td>0.08</td>
</tr>
<tr>
<td>Test Task Presentation (3, 57)</td>
<td>2.73+</td>
<td>0.13</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 57)</td>
<td>1.57</td>
<td>0.08</td>
</tr>
<tr>
<td>Electrode (4, 76)</td>
<td>10.46**</td>
<td>0.36</td>
</tr>
<tr>
<td>Condition x Electrode (4, 76)</td>
<td>&lt; 1</td>
<td>0.01</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (12, 228)</td>
<td>2.57+</td>
<td>0.12</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (12, 228)</td>
<td>&lt; 1</td>
<td>0.03</td>
</tr>
<tr>
<td>Hemisphere (1, 19)</td>
<td>26.88***</td>
<td>0.59</td>
</tr>
<tr>
<td>Condition x Hemisphere (1, 19)</td>
<td>&lt; 1</td>
<td>0.00</td>
</tr>
<tr>
<td>Test Task Presentation x Hemisphere (3, 57)</td>
<td>10.38***</td>
<td>0.35</td>
</tr>
<tr>
<td>Electrode x Hemisphere (4, 76)</td>
<td>&lt; 1</td>
<td>0.01</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Hemisphere (3, 57)</td>
<td>&lt; 1</td>
<td>0.01</td>
</tr>
<tr>
<td>Condition x Electrode x Hemisphere (4, 76)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode x Hemisphere (12, 228)</td>
<td>1.95+</td>
<td>0.09</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode x Hemisphere (12, 228)</td>
<td>&lt; 1</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Note: degrees of freedom are reported in parentheses in the left column. *p > .05 and < .10, **p < .05, ***p < .01, ****p < .001.
As shown in Table 3.2, there was a marginally significant effect of condition (nonword, word) across the midline sites \((p = .07)\) in the predicted direction (i.e., mean amplitude for the word condition was more positive than for the nonword condition), though condition did not interact with any other factors in any of the analyses. This may be related to the observed main effect of test task presentation order. Planned Bonferroni-corrected comparisons for the analysis at midline sites revealed that mean amplitude in the 300 to 600 millisecond time window at the Pre-Test did not differ significantly from that of Test 1 or Test 3, though it was marginally significant compared to Test 2 \((p = .06)\). Additionally, mean amplitude at Test 1 also differed from that at Test 2 \((p < .05)\), but not from that at Test 3. Finally, mean amplitude at Test 2 differed from that at Test 3 \((p < .05)\). In addition, for both the medial-lateral and lateral-lateral analyses, the mean amplitude of the overall waveforms was more negative for the left hemisphere relative to the right hemisphere, but the mean amplitude for the right hemisphere was much more variable across time than for the left hemisphere.

It is important to recall that Test 2 was a one-week retention task, administered one week after the participants' last exposure to the Dutch words, so some forgetting was certainly to be expected, which perhaps can account for why the monolinguals seemed to show an N400 to the Dutch words at Test 2, relative to the other tests. It seems as though the monolinguals were becoming more sensitive to the word status of the Dutch words they were learning, when comparing the mean amplitude of the words at Test 1 relative to the Pre-Test (learning across the first experimental session) and at Test 3 relative to Test 2 (learning across the
second experimental session). Unfortunately, the nonwords generally showed this same pattern of results, which may be due to the repetition of the nonwords within an experimental session and becoming more "word-like."

While the ERP data from the monolingual participants in Experiment 1a revealed main effects of test task presentation, which suggests differential processing over time, these changes did not depend on whether a stimulus was a nonword or a word. Except for a hint of a main effect of condition (nonword, word) in the midline analysis, no other main effects of condition or interactions with condition appear in the ERP analyses. It appears that in contrast to the results reported by McLaughlin et al. (2004)—evidence of learning in the ERP record in the absence of behavioral evidence—the monolingual data from Experiment 1a show evidence of learning in both the behavioral and ERP records. Although, as previously described, the two studies differed in a number of factors, it remains striking that evidence of learning could be observed in the behavioral record for the monolinguals without significant effects of condition or interactions between condition and electrode in the ERP analyses. Before discussing these results in more detail, however, data from the bilingual participants are reported.

3.1.4.2 Test task ERP Results: Bilinguals. After being subjected to offline post-processing as described in section 3.1.2.4, and after the exclusion of all data from participants with fewer than 13 trials remaining in any condition at any test task presentation after artifact rejection, 76.62% of trials remained in the monolingual data for the test task analyses. A 2 (condition: nonword, word) by 4 (test task presentations) repeated-measures ANOVA conducted on the percentage of
trials retained revealed no differences in condition \((F(1, 17) = 1.46, p = .24)\), no differences in task presentation \((F(3, 51 = 1.17, p = .33)\), and no interaction between condition and task presentation \((F(3, 57) < 1)\).

The average waveforms plotted for representative electrodes F3, FZ, F4, C3, CZ, C4, P3, PZ, and P4 for all four test task presentations (Pre-Test, Test 1, Test 2, and Test 3) can be found in Figures 3.8a, 3.8b, 3.8c, and 3.8d, respectively. Side-by-side plots for both monolinguals and bilinguals are plotted at representative electrode CZ for nonwords and words at each task presentation in Figure 3.9.

![Figure 3.8a](image_url)

Figure 3.8a. Grand average ERP waveforms for bilinguals at the Pre-Test for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms
represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.

Figure 3.8b. Grand average ERP waveforms for bilinguals at Test 1 for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 3.8c. Grand average ERP waveforms for bilinguals at Test 2 for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 3.8d. Grand average ERP waveforms for bilinguals at Test 3 for Experiment 1a. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. μV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Pre-Test

Test 1

Test 2

Test 3

Nonwords Words

Monolinguals Experiment 1a English Translations

Bilinguals Experiment 1a English Translations

μV 0

-150 0 300 700 ms
Figure 3.9. Grand average ERP waveforms for monolinguals and bilinguals from Experiment 1a at representative electrode site CZ for all test task presentations. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The y-axis represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.

As previously mentioned, the ERP data from the bilingual group was analyzed using three separate repeated-measures ANOVAs for midline electrodes, medial-lateral electrodes, and lateral-lateral electrodes. The results of these analyses are presented in Table 3.3.
Table 3.4. Bilingual ERP ANOVA results for Experiment 1a.

<table>
<thead>
<tr>
<th>Source</th>
<th>F-value</th>
<th>ηp²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 17)</td>
<td>20.52***</td>
<td>0.55</td>
</tr>
<tr>
<td>Test Task Presentation (3, 51)</td>
<td>10.69***</td>
<td>0.39</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 51)</td>
<td>1.62</td>
<td>0.09</td>
</tr>
<tr>
<td>Electrode (3, 51)</td>
<td>36.25***</td>
<td>0.68</td>
</tr>
<tr>
<td>Condition x Electrode (3, 51)</td>
<td>2.35</td>
<td>0.12</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (9, 153)</td>
<td>4.71***</td>
<td>0.22</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (9, 153)</td>
<td>1.81</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Medial-Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 17)</td>
<td>15.64**</td>
<td>0.48</td>
</tr>
<tr>
<td>Test Task Presentation (3, 51)</td>
<td>9.34***</td>
<td>0.36</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 51)</td>
<td>1.64</td>
<td>0.09</td>
</tr>
<tr>
<td>Electrode (6, 102)</td>
<td>11.63***</td>
<td>0.41</td>
</tr>
<tr>
<td>Condition x Electrode (6, 102)</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (18, 306)</td>
<td>2.95*</td>
<td>0.15</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (18, 306)</td>
<td>1.08</td>
<td>0.06</td>
</tr>
<tr>
<td>Hemisphere (1, 17)</td>
<td>8.25*</td>
<td>0.33</td>
</tr>
<tr>
<td>Condition x Hemisphere (1, 17)</td>
<td>5.79*</td>
<td>0.25</td>
</tr>
<tr>
<td>Test Task Presentation x Hemisphere (3, 51)</td>
<td>2.04</td>
<td>0.11</td>
</tr>
<tr>
<td>Electrode x Hemisphere (6, 102)</td>
<td>4.60**</td>
<td>0.21</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Hemisphere (3, 51)</td>
<td>1.16</td>
<td>0.06</td>
</tr>
<tr>
<td>Condition x Electrode x Hemisphere (6, 102)</td>
<td>2.07+</td>
<td>0.11</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode x Hemisphere (18, 306)</td>
<td>1.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode x Hemisphere (18, 306)</td>
<td>1.2</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Lateral-Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 17)</td>
<td>5.66*</td>
<td>0.25</td>
</tr>
<tr>
<td>Test Task Presentation (3, 51)</td>
<td>6.48**</td>
<td>0.28</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 51)</td>
<td>1.01</td>
<td>0.06</td>
</tr>
<tr>
<td>Electrode (4, 68)</td>
<td>6.85*</td>
<td>0.29</td>
</tr>
<tr>
<td>Condition x Electrode (4, 68)</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (12, 204)</td>
<td>5.59**</td>
<td>0.25</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (12, 204)</td>
<td>&lt; 1</td>
<td>0.05</td>
</tr>
<tr>
<td>Hemisphere (1, 17)</td>
<td>7.46*</td>
<td>0.31</td>
</tr>
<tr>
<td>Condition x Hemisphere (1, 17)</td>
<td>1.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Test Task Presentation x Hemisphere (3, 51)</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Electrode x Hemisphere (4, 68)</td>
<td>7.07**</td>
<td>0.29</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Hemisphere (3, 51)</td>
<td>1.3</td>
<td>0.07</td>
</tr>
<tr>
<td>Condition x Electrode x Hemisphere (4, 68)</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode x Hemisphere (12, 204)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode x Hemisphere (12, 204)</td>
<td>1</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Note: degrees of freedom are reported in parentheses in the left column. +p > .05 and < .10, *p < .05, **p < .01, ***p < .001
As shown in Table 3.3, and in contrast to the results for the monolinguals, a main effect of condition was observed for the midline, medial-lateral, and lateral-lateral analyses in the predicted direction: as the mean amplitude for nonword trials was more negative than for word trials. Additionally, in the medial-lateral analysis only, there was a significant condition by hemisphere interaction, such that the difference in mean amplitude between word and nonword trials was greater for left hemisphere electrodes than right hemisphere electrodes.

There was also a main effect of test task presentation order observed in all three analyses, and planned Bonferroni-corrected pairwise comparisons were performed for each analysis to determine which test task presentations differed from each other. For the midline analyses on the bilingual ERP data, the difference in mean amplitude between the Pre-Test and Tests 1 and 2 was marginally significant \((p = .05\) and \(p = .07\), respectively), but no difference was observed between the Pre-Test and Test 3. Test 1 differed significantly from Test 2 \((p < .001)\), and marginally so from Test 3 \((p = .10)\). Finally, Test 2 differed significantly from Test 4 \((p < .05)\). For the medial-lateral analysis, again, the difference in mean amplitude between the Pre-Test and Tests 1 and 2 was marginally significant \((p = .06\) and \(p = .07\), respectively), but no difference was observed between the Pre-Test and Test 3. Test 1 differed from Test 2 \((p < .001)\), but not from Test 3, and Test 3 differed from Test 4 \((p < .05)\). For each of the four test task presentations, the overall mean amplitude was 1.56 for the Pre-Test, 1.94 for Test 1, 0.72 for Test 2, and 1.40 for Test 3. Finally, an interaction between test task presentation order and
electrode was observed for all three analyses, but seems largely driven by little variation in amplitude at the occipital sites (midline, medial-lateral, and lateral-lateral analyses) and the temporal-parietal sites (lateral-lateral analysis only).

3.1.5. Experiment 1a Discussion. Several results differed in the ERP analyses for the monolinguals and bilinguals. First and most importantly, the effect of condition (i.e., more negative mean amplitude for nonwords than for words) was much more robust for the bilinguals across all analyses, whereas for the monolinguals, only a marginally significant main effect of condition was found in the midline analysis. However, behavioral evidence showed clear effects of learning by both groups, both within sessions and across them. These results suggest two possibilities. First, it may be the case that the bilinguals are faster learners, and that the evidence of learning detected in the electrophysiological record at these first stages of learning are necessary phases through which the monolingual learners—had they continued to study Dutch vocabulary beyond the three testing phases—would eventually pass through. However, this account would have difficulty addressing how the monolinguals were learning any vocabulary at all, given the lack of differentiation by the brain to the nonwords and words. It seems much more likely that the context of learning in which these participants engaged was different for the two groups not due to experimental manipulation, but due to the language processing differences the two groups were engaging in during the study and test tasks.
As Bogulski and Kroll (under review) argued, the context of learning foreign language vocabulary via L1 translations is a very different experience for a bilingual than for a monolingual. In the bilingual case, the learner has had experience learning to inhibit the L1 in order to attend to and produce the L2. A monolingual learner certainly has inhibitory experience in a language context (e.g., experiencing interference in word retrieval due to a similar sounding word), but it does not begin to approach the amount of inhibitory processing the bilingual must exercise in order to successfully function within an L2. Hence, in the context of foreign language vocabulary learning, bilinguals learning via L1 translations may be experiencing added interference from the L1 during the study task, as the language learning context may activate L1-inhibitory processes. This inhibitory hypothesis proposed by Bogulski and Kroll could then be extended to the results of Experiment 1a, implying that learning foreign language vocabulary via L1 translations as a bilingual has early consequences for language processing in the newly-acquired language. Furthermore, the ERP results of Experiment 1a support the idea that this context of learning changes how the brain responds to linguistic stimuli in an unfamiliar language the bilingual is trying to learn. Hence, the combined results of Experiment 1a suggest that the bilingual advantage in foreign language vocabulary learning extends to a level of neurophysiological sensitivity in a laboratory learning environment. This evidence then supports the idea that the bilingual advantage in foreign language vocabulary learning is not simply an explicit strategy difference employed by the two groups, but rather is a manifestation of more general
consequences to neural networks. Whether these consequences are similar for all bilinguals regardless of learning context is a question that was investigated in Experiment 1b.
CHAPTER 4: EXPERIMENT 1B – RE-EXAMINING THE BILINGUAL ADVANTAGE IN FOREIGN LANGUAGE VOCABULARY LEARNING VIA L2 TRANSLATIONS USING ERPS

The purpose of Experiment 1b was twofold: to further examine bilingual brain responses during a foreign language vocabulary learning task, and to test the inhibitory hypothesis proposed by Bogulski and Kroll (under review). The results of both Bogulski and Kroll and Experiment 1a suggest that bilinguals are advantaged relative to monolinguals in learning foreign language vocabulary. Moreover, the results of Experiment 1a further support the idea that this advantage is reflected in electrophysiological measures of the brain.

4.1 Experiment 1b

4.1.1 Method

4.1.1.1 Participants. Fifteen English-Spanish bilinguals were recruited for Experiment 1b. Data from one participant was lost due to experimenter error, and data from one additional participant were discarded due to a lack of a sufficient number of trials after artifact rejection per condition for inclusion in the ERP analysis. Finally, data from one bilingual participant were discarded due to this person speaking Spanish as a home language. The results presented for Experiment 1b reflect data from the remaining 12 bilingual participants. Like Experiment 1b, the English-Spanish bilinguals recruited for this study were predominantly late learners
of Spanish (L2 Onset $M = 14.18$, $SD = 1.66$; range 11 to 17, 11 participants reporting).

The bilinguals in Experiment 1b were recruited to match the bilinguals tested in Experiment 1a as closely as possible in age, L1 proficiency, L2 proficiency, and age of L2 onset. Indeed, one-way ANOVAs performed for factors of number of days between sessions 1 and 2, age, and native language proficiency self-ratings across all three groups (i.e., monolinguals and bilinguals from Experiment 1a and bilinguals from Experiment 1b) revealed no significant differences between the groups ($F < 1$ for days between sessions; $F < 1$ for age; $F = 1.78$, $p = .18$ for native language proficiency self-rating). An additional one-way ANOVA performed for the three groups for L2 proficiency self-ratings revealed a main effect ($F = 30.76$, $p < .001$) that posthoc tests (Tukey’s HSD) further revealed was driven entirely by the low ratings of the monolinguals, as the two bilingual groups did not differ from one another ($t = 7.38$, $p < .001$ for the comparison between monolinguals and bilinguals in Experiment 1a, $t = 5.51$, $p < .001$ for the comparison between monolinguals and the bilinguals tested in Experiment 1b). Additionally, age of L2 onset did differ marginally across the three groups ($F = 2.60$, $p = .09$), but critically, posthoc tests (Tukey’s HSD) revealed that the two bilingual groups did not differ from one another ($p = .21$).

Like Experiment 1a, all participants tested in Experiment 1b were recruited from either the campus or the outlying areas surrounding the Pennsylvania State University, and were compensated $10 per hour for their participation. Again, like
Experiment 1a, Experiment 1b consisted of two sessions approximately one week apart (range: 5 to 10 days, mean = 7.1 days), and participants were compensated for their time at the end of each session.

4.1.2 Materials and Design. The 90 Dutch words taught to the bilingual participants in Experiment 1b were identical to those used in Experiment 1a (see Appendix. However, as the nature of the research question being asked in Experiment 1b concerned learning foreign language vocabulary via L2 translations, all 90 English translations of the Dutch words were translated into Spanish for the study task. As the test task presented only the 90 Dutch words and the 90 matched Dutch-like nonwords, the test task was identical to that of Experiment 1a. The experiment was identical to Experiment 1a other than the pairing of Dutch words to their Spanish (L2) translations.

All other tasks and materials used in Experiment 1b were identical to those in Experiment 1a.

4.1.3 Procedure. The procedure of Experiment 1b was identical to that of Experiment 1a, save for Spanish being the language of translation during the study task. Hence, the English-Spanish bilinguals in Experiment 1b were asked to read and produce the Spanish translation of the presented Dutch word during the study task. Additionally, at the end of the second and final experimental session, participants were asked to complete a translation elicitation task (Kroll & Stewart, 1994) using paper and pencil in order to assess the participants' knowledge of the Spanish items used during the experiment. Each of the 90 Spanish words that the bilinguals had
been exposed to throughout the experiment as correct translations for the Dutch vocabulary items were presented, and the participants’ task was to provide the correct English translation for each Spanish word.

**4.1.4 EEG acquisition and offline processing.** The method of EEG acquisition and offline processing in Experiment 1b was identical to that in Experiment 1a.

**4.1.2 Results: Behavioral Data.** The goal of Experiment 1b was to determine whether the bilingual advantage in foreign language vocabulary learning would extend to a bilingual group learning the new vocabulary via L2 translations. As previously reviewed, Bogulski and Kroll (under review) reported no bilingual advantage for two bilingual groups learning foreign language vocabulary via L2 translations. Experiment 1b addresses two issues raised by those results. First, Experiment 1b tested another group of English-Spanish bilinguals, which is a group that has consistently demonstrated the bilingual advantage in foreign language vocabulary learning, both in previous work and in Experiment 1a (Bogulski & Kroll, under review; Kaushanskaya & Marian, 2009a, 2009b). Second, Experiment 1b is the first to record electrophysiological brain activity using ERPs to examine the bilingual advantage in foreign language vocabulary learning via L2 translations. Like Experiment 1a, the results of Experiment 1b are presented and discussed for the behavioral data first, and then re-examined in the context of the ERP evidence. In order to directly compare the results of Experiment 1b to those from Experiment 1a, the figures and analyses in Experiment 1b will include not only the bilingual
participants learning via L2 translations (bilingual L2 learners), but also the monolingual and bilingual participants (bilingual L1 learners) from Experiment 1a.

Participant descriptive statistics across the three groups tested, including number of participants per group, mean age, number of days between sessions, mean age of L2 onset, mean L1 self-ratings, mean L2 self-ratings, and Operation-Span score, can be found in Table 4.1 below.

Table 4.1. Mean age, number of days between sessions, age of L2 onset, L1 self-ratings, L2 self-ratings, and Operation-Span score for the monolinguals, bilingual L1 learners, and bilingual L2 learners test in Experiments 1a and 1b (standard deviations in parentheses, significant differences between groups are denoted with *; marginal significance is denoted with †).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age (SD)</th>
<th>Days Between Session 1 &amp; 2 (SD)</th>
<th>Age of L2 Onset (SD)</th>
<th>L1 Self-Rating (SD)</th>
<th>L2 Self-Rating* (SD)</th>
<th>Operation-Span Score* (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolinguals</td>
<td>20</td>
<td>22.0 (3.7)</td>
<td>7.1 (0.4)</td>
<td>14.3 (2.1)</td>
<td>9.8 (0.3)</td>
<td>2.9 (0.2)</td>
<td>44.8 (8.8)</td>
</tr>
<tr>
<td>Bilingual L1 Learners</td>
<td>18</td>
<td>21.8 (2.7)</td>
<td>6.8 (0.4)</td>
<td>12.5 (3.3)</td>
<td>9.6 (0.3)</td>
<td>7.1 (0.5)</td>
<td>49.9 (6.4)</td>
</tr>
<tr>
<td>Bilingual L2 Learners</td>
<td>12</td>
<td>23.5 (6.2)</td>
<td>7.1 (1.2)</td>
<td>14.2 (1.7)</td>
<td>9.9 (0.7)</td>
<td>6.4 (1.6)</td>
<td>49.5 (8.3)</td>
</tr>
</tbody>
</table>

One-way ANOVAs were conducted on all of the variables for all three groups depicted in Table 4.1 except for Operation Span score, which was limited to just the bilingual L2 learners and the monolinguals, as no differences were predicted between the two bilinguals groups for working memory span. No differences existed among the groups for age ($F < 1$), the number of days between Sessions 1 and 2 ($F < 1$), or L1 self-ratings ($F(2, 47) = 1.78, p = .18$). The difference in age of L2 onset was
marginally significant \((F(2, 47) = 2.60, p = .09)\), as both bilingual groups overall had their first exposure to a second language earlier than did the monolinguals (though planned post-hoc comparisons did not reveal any significant differences between the groups). Additionally, a main effect of L2 self-ratings was observed \((F(2, 47) = 30.76, p< .001)\), and planned post-hoc comparisons (Tukey’s HSD) revealed that both bilingual groups rated themselves as much more proficient in an L2 than did the monolinguals \((p < .05)\). Finally, the main effect of Operation Span score was not significant in the one-way ANOVA comparing monolinguals and bilingual L2 learners, though the bilingual L2 learners trended toward outperforming the monolinguals on the Operation Span task \((F(1, 30) = 2.24, p = .15)\).

**4.1.2.1 Flanker task.** To analyze the data from the flanker task, comparisons were limited to those between the monolinguals from Experiment 1a and the bilingual L2 learners tested in Experiment 1b. The accuracy rates for both the monolinguals and the bilingual L2 learners on all trial types in the flanker task was very high (all group averages per condition were at least 90%). Due to these high levels of accuracy, the accuracy data will not be analyzed any further.

The reaction time data were analyzed using five separate ANOVAs. First, the Flanker Effect was calculated as the speed of response to incongruent trials relative to congruent trials, in an experimental block containing only those two types of trials. A one-way ANOVA revealed no differences between the two groups in the Flanker Effect \((F(1, 30) = 1.36, p = .25)\). Next, speed of performance for each trial type (i.e., incongruent, congruent, go, and control) in the blocked condition was
compared using a 2 (group: monolinguals, bilingual L2 learners) by 4 (trial type: incongruent, congruent, go, and control) mixed-factor ANOVA. The results showed no main effect of group, nor a group by trial type interaction (both \(F_s < 1\)), but did reveal a main effect of trial type (\(F(3, 90) = 118.55, p < .001, \eta_p^2 = .80\)). Planned Bonferroni-corrected pairwise comparisons revealed that control trials were faster than all other trial types (all \(p_s < .001\)), followed by congruent trials, which were faster than both go and incongruent trials (both \(p\) values < .01). No difference between go and incongruent trials was found.

Next, a 2 (group: monolinguals, bilingual L2 learners) by 3 (trial type: incongruent, congruent, go) mixed-factor ANOVA was conducted on the reaction time data for those trials that were in the mixed block, which contained incongruent, congruent, go, and no-go trials. The results again demonstrated no main effect of group (\(F < 1\)), nor a group by trial type interaction (\(F < 1\)). A main effect of trial type did emerge (\(F(2, 60) = 36.20, p < .001, \eta_p^2 = .57\)), and as planned Bonferroni-corrected pairwise comparisons revealed, the congruent trials were significantly faster than go trials, which were significantly faster than incongruent trials (all \(p\) values < .01).

A fourth ANOVA was conducted on the mixing cost data, which were difference scores calculated by subtracting reaction times for incongruent, congruent, and go trials in the blocked condition from the reaction times for those same trials in a mixed block. This 2 (group: monolinguals, bilingual L2 learners) by 3 (trial type: incongruent, congruent, go) mixed-factor ANOVA revealed no main effect
of group ($F < 1$), and the group by trial type interaction no longer approached
significance ($F < 1, \eta^2_p = .02$), as the two groups experienced similar costs for
congruent, incongruent and go trials.

Finally, an additional 2 (group: monolinguals, bilingual L2 learners) by 3
(trial type: incongruent, congruent, go) mixed-factor ANOVA was conducted on the
relative costs for the incongruent, congruent, and go trials. Relative costs are
calculated by comparing the speed of response for each of the critical trials to the
control trials to assess how much of an additional processing cost each type of trial
imposes. This analysis revealed no main effect of group nor a group by trial type
interaction (both $F$s < 1), though a main effect of trial type was again significant
($F(2, 60) = 14.12, p < .001, \eta^2_p = .32$). Planned Bonferroni-corrected pairwise
comparisons revealed that costs to both incongruent and go trials was greater than
the cost to congruent trials (both $p$ values < .001), but there was no difference in
costs to incongruent and go trials ($p = 1.0$).

Unlike the balanced bilinguals tested by Luk (2008), and in contrast to the
results reported for the flanker task in Experiment 1a, a bilingual advantage in the
mixing cost to go trials was not observed when in Experiment 1b. In fact, no group
differences were observed in the analyses of the flanker task data in Experiment 1b.
Even though the bilingual L2 learners were drawn from the same population as the
bilingual L1 learners tested in Experiment 1a, the sample sizes in both experiments
was fairly small. Hence, the differences between the results of the flanker analyses
in Experiment 1a and Experiment 1b will be further discussed in the General Discussion.

4.1.2.2 Word naming at study: Behavioral data. Overall, the English-Spanish bilinguals tested in Experiment 1b correctly produced the target Spanish translation during the study task 99.7%, 99.9%, and 100% of the time, for each of the three presentations of the study task, respectively. The remaining trials were either trials named incorrectly (misread/mispronounced, or said the Dutch word aloud), or not named at all. Despite the fact that Spanish was the L2 for this group, the Spanish word was presented on the computer screen during this task, making the production portion of this task essentially L2 word naming, which accounts for the highly accurate performance in L2 on this task. Due to the high level of accuracy all presentations of the study task, the accuracy data will not be analyzed any further.

For the naming latency analyses, only trials identified as named correctly with accurate voice-key triggering were included. These naming latencies for both the monolingual and bilingual groups for all three presentations of the study task are presented in Figure 4.1.
A 3 (group: monolinguals, bilingual L1 learners, bilingual L2 learners) by 3 (study task presentations) mixed-factor ANOVA revealed no main effect of group ($F(2, 46) = 1.15, p = .33, \eta_p^2 = .05$) or of presentation order ($F(2, 92) = 1.39, p = .26, \eta_p^2 = .03$). It did, however, reveal a group by presentation order interaction ($F(4, 84) = 2.91, p < .05, \eta_p^2 = .11$). As the planned comparisons for the group and time factors revealed no significant differences, the interaction seems to be driven largely by the bilingual L2 learners being slowest to name translations at study for the first presentation of the study task, but naming faster than the bilingual L1 learners (though still slower than the monolingual learners) to name these same translations at the second and third presentations. As the monolingual and bilingual L1 learners were naming in English, and the bilingual L2 learners were naming in Spanish, these comparatively fast word naming latencies for the bilingual L2 learners are quite striking. Also
noteworthy is the fact that the second presentation of the study task occurred within the same experimental session as the first (Session 1), whereas the third presentation of the study task occurred in Session 2. Hence, the initially slow (albeit variable) Spanish naming latencies during the Dutch learning task might reasonably be hypothesized to be due to slower L2 lexical access. Much more remarkable is the dramatic decrease in L2 naming latencies, both within the same session and one week later.

Much like the L1-dominant Chinese-English bilinguals in Bogulski and Kroll (under review), who were learning Dutch vocabulary via English (L2) translations, the bilinguals in Experiment 1b were much faster to name L2 translations aloud during the study task than would be expected if they were engaged in similar processes and strategies as the bilinguals learning via L1 translations as well as suffering from a deficit due to L2 lexical access. Hence, the inhibitory hypothesis would predict that the reduced interference the bilingual L2 learners exhibit during the study task would not lead to the presence of a bilingual advantage in foreign language vocabulary learning at the test task.

4.1.2.3 Recognition test task: Behavioral data. The data from all four test tasks for Experiments 1a and 1b are presented in Figures 4.2 (accuracy) and 4.3 (reaction times).
A 2 (group: monolinguals, bilinguals) by 4 (test task presentations) mixed-factor ANOVA performed on the $d'$ data revealed a marginally significant main effect of group ($F(2, 47) = 2.36, p = .11, \eta^2_p = .09$). While planned comparisons between the groups (Tukey’s HSD) did not reveal any significant differences between any two particular groups, the comparison that approached significance most closely was that between the monolinguals and the bilingual L1 learners from Experiment 1a ($p = .12$). Additionally, a main effect of task presentation order emerged ($F(3, 141) = 115.48, p < .001, \eta^2_p = .71$). Planned tests were performed to determine which of the four presentations of the test task differed between each of the others. Planned Bonferroni-corrected pairwise comparisons revealed that participants were significantly more accurate at all critical tests than at the Pre-Test (Pre-Test to Test 1: $t = 12.21, p < .001$; Pre-Test to Test 2: $t = 10.08, p < .001$; Pre-Test to Test 3: $t =$
12.46, \( p < .001 \). Additional planned Bonferroni-corrected pairwise comparison further revealed that participant performance at Test 1 was more accurate than at Test 2 (\( t = 8.96, p < .001 \)), but not at Test 3 (\( t < 1, p = 1.0 \)). Further, performance at Test 3 was more accurate than performance at Test 2 (\( t = 8.07, p < .001 \)). Finally, and perhaps most critically, in addition to the two main effects, a marginally significant group by test task presentation interaction emerged (\( F(6, 141) = 2.60, p = .07, \eta^2_p = .09 \)). Given that performance between the three groups was similar at the pre-test, the interaction seems to be driven by the two bilingual groups (bilingual L1 learners and bilingual L2 learners) outperforming the monolingual group at all three critical tests, though the differences were largest at Test 1 and Test 3.

Contrary to the predictions made by the inhibitory hypothesis proposed by Bogulski and Kroll (under review), both bilingual L1 and bilingual L2 learners exhibited an advantage in foreign language vocabulary learning relative to the monolingual learners for immediate tests. However, a benefit for learning foreign language vocabulary via L1 translations did emerge, as the bilingual L1 learners outperformed both the monolinguals and the bilingual L2 learners at the retention task at the beginning of Session 2, approximately one week after they had last been exposed to the vocabulary.
Figure 4.3a: Test task reaction times on nonword trials for Experiments 1a and 1b.

Figure 4.3b: Test task reaction times on word trials for Experiments 1a and 1b.
A 3 (group: monolinguals, bilingual L1 learners, bilingual L2 learners) by 4 (test task presentations) by 2 (condition: nonword, word) mixed-factor ANOVA on test task reaction times revealed no main effect of group \( (F(2, 46) = 1.53, p = .23, \eta^2_p = .06) \), but a marginally significant group by task presentation interaction emerged \( (F(6, 138) = 3.67, p < .05, \eta^2_p = .14) \), which seems to be driven by the bilingual L1 learners performing the task more slowly than the monolinguals and the bilingual L2 learners, but this difference being present only for the critical tests (Tests 1, 2, and 3) and not the Pre-Test. Additionally, a main effect of condition emerged, as nonword trials were slower than word trials overall \( (F(1, 46) = 5.56, p < .05) \). This effect of condition was qualified by a significant task presentation by condition interaction \( (F(3, 138) = 15.30, p < .001) \), such that the difference in the speed of responses to nonword trials was largest for Tests 1 and 3, and smaller or non-existent for the Pre-Test and Test 2. Finally, a marginally significant three-way interaction between group, task presentation, and condition emerged \( (F(6, 138) = 1.93, p = .09, \eta^2_p = .08) \), as the differences between the groups was greater for the nonwords than for the words (the bilingual L1 learners being the slowest to perform the task), but this difference was largely absent in the Pre-Test, as well as during Test 3. Finally, a main effect of task presentation order emerged \( (F(3, 138) = 25.09, p < .001, \eta^2_p = .35) \). Planned Bonferroni-corrected pairwise comparisons revealed that reaction times for the Pre-Test were slower than at all three critical tests \( (Test 1, t = 4.09, p < .01; Test 2, t = 4.16, p < .01; Test 3, t = 6.33, p < .001) \).
Additionally, performance at Test 3 (the final test) was faster than at both Tests 1 \((t = 5.25, p < .001)\) and 2 \((t = 6.16, p < .001)\). No other effects approached significance.

While overall, participants generally performed the test task more quickly over time, the bilingual L2 learners performed the task much more like the monolingual learners. The bilingual L1 learners performed the task more slowly and did not speed up over the test task presentations as much as the monolinguals and the bilingual L2 learners.

It is unsurprising that performance at Test 1 was faster than at the Pre-Test, given that by Test 1, participants had previously been given two opportunities to study the Dutch words, in addition to the Pre-Test, which may have familiarized them to the words. Additionally, given that approximately one week elapsed between test 2 and test 3, it is also unsurprising that participants were slower and less accurate at test 3 relative to their performance at test 2.

The most striking result from the test task data is that the bilingual L2 learners performed as quickly and as accurately as they did, given that they had been studying the words via Spanish (L2) translations. Given the added difficulty of learning via L2 translations, the bilingual L2 learners might have been expected to perform even more slowly on the study task given that they were naming in their L2. These faster naming latencies for L2 translations during the study task were predicted, however, given the results from the bilingual L2 learners reported by Bogulski and Kroll (under review). However, the bilingual L2 learners reported in Experiment 1b, unlike the bilingual L2 learners reported by Bogulski and Kroll, not
only demonstrated a bilingual advantage in foreign language vocabulary learning, they were faster at the test task as well. Perhaps the most critical difference between the bilingual L2 learners tested in Experiment 1b and the bilingual L2 learners tested by Bogulski and Kroll was that the L2 learners in Experiment 1b were tested in their dominant language environment. The role of L2 immersion has previously been identified as having profound consequences for L1 language processing (e.g., Dussias & Sagarra, 2007; Linck, Kroll, & Sunderman, 2009). However, the role of L1 immersion for foreign language learning has not yet been empirically contrasted with L2 immersion. The role of L1 immersion in foreign language vocabulary learning will be further examined in the General Discussion. At present, the ERP data from the bilingual L2 learners will be examined and compared to that from the monolingual and bilingual L1 learners to investigate the neurophysiological consequences of foreign language vocabulary learning.

4.1.2 Results: ERP Data.

4.1.2.1 Test task ERP Results. After being subjected to offline post-processing as described in Chapter 3, and after the exclusion of all data from participants with fewer than 13 trials remaining in any condition at any test task presentation after artifact rejection, 75.42% of trials remained in the data for the test task analyses. A 2 (condition: nonword, word) by 4 (test task presentations) within-subjects ANOVA conducted on the percentage of trials retained revealed no differences in condition ($F(1, 11) = 1.49, p = .25, \eta_p^2 = .12$), no significant effect of time ($F(3, 33) < 1$), although the interaction between condition and task
presentation was marginally significant ($F(3, 33) = 2.83, p = .053, \eta^2 = .21$).

Nevertheless, the percentage of trials kept per task presentation, per condition always ranged between 71% and 77%. As differences between conditions across test task presentations in blinking or muscle artifacts are not of theoretical interest, it is not discussed further.

The average waveforms plotted for representative electrodes F3, FZ, F4, C3, CZ, C4, P3, PZ, and P4 for all four test task presentations (Pre-Test, Test 1, Test 2, and Test 3) can be found in Figures 4.4a, 4.4b, 4.4c, and 4.4d, respectively. To allow for a visual comparison to the monolinguals and bilinguals from Experiment 1a, representative electrode CZ is plotted for all 3 groups across all four test task presentations in Figure 4.5.
Figure 4.4a. Grand average ERP waveforms for bilinguals at the Pre-Test for Experiment 1b. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. μV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.

Figure 4.4b. Grand average ERP waveforms for bilinguals at Test 1 for Experiment 1b. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. μV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 4.4c. Grand average ERP waveforms for bilinguals at Test 2 for Experiment 1b. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 4.4d. Grand average ERP waveforms for bilinguals at Test 3 for Experiment 1b. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed green line found at 0 ms represents the onset of the stimulus during the Pre-Test. μV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.
Figure 4.5. Grand average ERP waveforms for monolinguals and bilinguals from Experiment 1a and bilinguals from Experiment 1b at representative electrode site CZ for all test task presentations. Nonword waveforms are plotted in black, and word waveforms are plotted in red. The vertical dashed line found at 0 ms represents the onset of the stimulus during the Pre-Test. uV are plotted on the y-axis, and milliseconds are plotted on the x-axis. Negative is plotted up.

As previously mentioned, the ERP data from the bilingual group tested in Experiment 1b was analyzed using three separate repeated-measures ANOVAs for midline electrodes, medial-lateral electrodes, and lateral-lateral electrodes. The results of these analyses are presented in Table 4.2.
Table 4.2. Bilingual ERP ANOVA results for Experiment 1b.

<table>
<thead>
<tr>
<th></th>
<th>F-value</th>
<th>( \eta^2_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 48)</td>
<td>16.34***</td>
<td>0.62</td>
</tr>
<tr>
<td>Test Task Presentation (3, 30)</td>
<td>8.49***</td>
<td>0.46</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 30)</td>
<td>1.92</td>
<td>0.16</td>
</tr>
<tr>
<td>Electrode (3, 30)</td>
<td>7.41**</td>
<td>0.43</td>
</tr>
<tr>
<td>Condition x Electrode (3, 30)</td>
<td>3.95*</td>
<td>0.28</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (9, 90)</td>
<td>2.41*</td>
<td>0.19</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (9, 90)</td>
<td>1.95*</td>
<td>0.16</td>
</tr>
<tr>
<td>Medial-Lateral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 10)</td>
<td>10.92**</td>
<td>0.52</td>
</tr>
<tr>
<td>Test Task Presentation (3, 30)</td>
<td>8.16*</td>
<td>0.45</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 30)</td>
<td>2.19*</td>
<td>0.18</td>
</tr>
<tr>
<td>Electrode (6, 60)</td>
<td>2.43*</td>
<td>0.20</td>
</tr>
<tr>
<td>Condition x Electrode (6, 60)</td>
<td>5.47*</td>
<td>0.35</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (18, 180)</td>
<td>1.95*</td>
<td>0.16</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (18, 180)</td>
<td>&lt; 1</td>
<td>0.08</td>
</tr>
<tr>
<td>Hemisphere (1, 10)</td>
<td>8.16*</td>
<td>0.45</td>
</tr>
<tr>
<td>Condition x Hemisphere (1, 10)</td>
<td>6.95*</td>
<td>0.41</td>
</tr>
<tr>
<td>Test Task Presentation x Hemisphere (3, 30)</td>
<td>2.01</td>
<td>0.17</td>
</tr>
<tr>
<td>Electrode x Hemisphere (6, 60)</td>
<td>1.52</td>
<td>0.13</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Hemisphere (3, 30)</td>
<td>&lt; 1</td>
<td>0.09</td>
</tr>
<tr>
<td>Condition x Electrode x Hemisphere (6, 60)</td>
<td>3.38**</td>
<td>0.25</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode x Hemisphere (18, 180)</td>
<td>1.96*</td>
<td>0.16</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode x Hemisphere (18, 180)</td>
<td>&lt; 1</td>
<td>0.07</td>
</tr>
<tr>
<td>Lateral-Lateral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (1, 10)</td>
<td>7.77*</td>
<td>0.44</td>
</tr>
<tr>
<td>Test Task Presentation (3, 30)</td>
<td>1.52</td>
<td>0.13</td>
</tr>
<tr>
<td>Condition x Test Task Presentation (3, 30)</td>
<td>2.36*</td>
<td>0.19</td>
</tr>
<tr>
<td>Electrode (4, 40)</td>
<td>&lt; 1</td>
<td>0.06</td>
</tr>
<tr>
<td>Condition x Electrode (4, 40)</td>
<td>3.84*</td>
<td>0.28</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode (12, 120)</td>
<td>2.77*</td>
<td>0.22</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode (12, 120)</td>
<td>&lt; 1</td>
<td>0.09</td>
</tr>
<tr>
<td>Hemisphere (1, 10)</td>
<td>15.76**</td>
<td>0.61</td>
</tr>
<tr>
<td>Condition x Hemisphere (1, 10)</td>
<td>&lt; 1</td>
<td>0.00</td>
</tr>
<tr>
<td>Test Task Presentation x Hemisphere (3, 30)</td>
<td>4.02*</td>
<td>0.29</td>
</tr>
<tr>
<td>Electrode x Hemisphere (4, 40)</td>
<td>&lt; 1</td>
<td>0.05</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Hemisphere (3, 30)</td>
<td>&lt; 1</td>
<td>0.04</td>
</tr>
<tr>
<td>Condition x Electrode x Hemisphere (4, 40)</td>
<td>&lt; 1</td>
<td>0.02</td>
</tr>
<tr>
<td>Test Task Presentation x Electrode x Hemisphere (12, 120)</td>
<td>1.1</td>
<td>0.10</td>
</tr>
<tr>
<td>Condition x Test Task Presentation x Electrode x Hemisphere (12, 120)</td>
<td>&lt; 1</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note: degrees of freedom are reported in parentheses in the left column. *p > .05 and < .13, **p < .05, ***p < .01, ****p < .001
As shown in Table 4.2, like the bilinguals and unlike the monolinguals in Experiment 1a, a main effect of condition was observed for the midline, medial-lateral, and lateral-lateral analyses in the predicted direction: as the mean amplitude for nonword trials was more negative than for word trials. Additionally, and like the results for the bilinguals in Experiment 1a, in the medial-lateral analysis only, there was a significant condition by hemisphere interaction, such that the difference in mean amplitude between word and nonword trials was slightly greater for left hemisphere electrodes than right hemisphere electrodes. A condition by test task presentation interaction was also marginally significant for both the medial-lateral and lateral-lateral analyses, as mean amplitude differences between nonwords (which tended to be more negative) and words were either diminished (medial-lateral) or reversed (lateral-lateral) for the Pre-Test. There was also a marginally significant condition by test task presentation by electrode interaction for the midline analysis only, as the differences in mean amplitude for nonwords and words were largest was present over central and parietal sites, and for Tests 1 and 3. Lastly, a condition by electrode by hemisphere interaction was observed in the medial-lateral analysis only, as larger differences in mean amplitude for nonwords relative to words were observed over central and parietal sites, which was especially the case for left hemisphere sites.

There was also a main effect of test task presentation order observed in the midline and medial-lateral analyses, and planned Bonferroni-corrected pairwise comparisons were performed for each analysis to determine which test task
presentations differed from each other. For the midline analysis on the bilingual ERP data for Experiment 1b, the difference in mean amplitude between the Pre-Test and Tests 1 was marginally significant \( (p = .06) \), and the mean amplitude at Test 2 was more negative than at Test 1 or Test 3. No other comparisons approach significance in the midline analysis. For the medial-lateral analysis, mean amplitude at the Pre-Test was slightly more negative than at Test 1 \( (p = .06) \), as well as at Test 2 relative to Test 1 \( (p = .07) \). Test 2 mean amplitude was also marginally more negative than at Test 3 \( (p = .10) \). Additionally, a task presentation by electrode by hemisphere interaction was marginally significant in the medial-lateral analysis, as mean amplitude was generally more negative at Tests 1 and 3, but this was especially true for frontal and right hemisphere.

A main effect of electrode site was significant in the midline analysis, as mean amplitude at FZ was more negative than at CZ, PZ, or OZ, according to planned Bonferroni-corrected pairwise comparisons \( (p < .05 \text{ for all three comparisons}) \). The main effect of electrode site was also marginally significant in the medial-lateral analysis, though planned Bonferroni-corrected pairwise comparisons revealed that the only comparison that approached significance was for frontal electrode sites (F3 and F4), which were more negative than at frontopolar sites (FP1 and FP2).

Additionally, a condition by electrode interaction was significant for the midline and medial-lateral analyses, and was marginally significant for the lateral-lateral analyses. For the midline analysis, the difference in mean amplitude between nonwords (more negative) and words (more positive) was largest for CZ and PZ,
and smaller for FZ and OZ, consistent with the typical topography of the N400. For the medial-lateral analysis, the mean amplitude was again greatest for nonwords relative to words for central and parietal sites, and smaller for frontal sites. Lastly, for the lateral-lateral analysis, the mean amplitude for nonword trials relative to word trials was greatest for temporal and parietal sites, and smaller again for frontal sites.

A main effect of hemisphere was found in both the medial-lateral and lateral-lateral analyses, as left hemisphere sites exhibited more negative mean amplitude than did right hemisphere sites. However, this main effect should be qualified for the lateral-lateral case, as there were greater mean amplitude differences between test task presentations for right hemisphere sites than for left.

The interaction between task presentation and electrode was significant in the lateral-lateral analysis, and marginally significant in both the midline and medial-lateral analyses. In the lateral-lateral analysis, the mean amplitude was most negative for the Pre-Test and Test 3, and this was especially the case for frontal and temporal sites. This same pattern largely held for the medial-lateral and midline analyses, although parietal sites also exhibited this pattern to some extent, though not occipital sites.

Finally, an interaction between test task presentation order and electrode was observed for all three analyses, but seems largely driven by little variation in amplitude at the occipital sites (midline, medial-lateral, and lateral-lateral analyses) and the temporal-parietal sites (lateral-lateral analysis only).
4.1.2.1.1. Discussion of Bilingual ERP results. Although fewer subjects participated in Experiment 1b, the ERP results in Experiment 1b are therefore all the more striking. Like the bilinguals in Experiment 1a, effects of condition were obtained in all three analyses. However, unlike the results from the bilinguals in Experiment 1a, the bilinguals in Experiment 1b exhibited the greatest effect of condition (i.e., enhanced negativity to nonwords relative to words) for central and parietal sites, which is consistent with typical N400 elicitations. Moreover, for these bilinguals, though statistical significance was marginal, the sensitivity to nonwords relative to words was greatest for Tests 1 and 3 (the tests at the end of an experimental session that included additional study and testing opportunities), and smallest for the Pre-Test.

What makes these results especially surprising is the fact that the behavioral performance—as indexed by $d'$—of these bilinguals across Tests 1, 2, and 3 was similar to that of the bilinguals in Experiment 1a. In contrast to predictions, bilinguals learning foreign language vocabulary words via L2 translations also demonstrated a bilingual advantage relative to the monolinguals, and also exhibited greater sensitivity to the Dutch-like nonwords relative to the real Dutch words across the critical test task presentations relative to both the monolinguals and the bilinguals from Experiment 1a.

These results fail to support the inhibitory hypothesis proposed by Bogulski and Kroll (under review), who argued that the bilingual advantage in foreign
language vocabulary learning requires a context of learning in which interference from the dominant language ultimately predicts better learning at test. It is difficult to surmise what the bilinguals in Experiment 1b might be doing during the study task, since they were actually faster to perform it than were the bilinguals in Experiment 1a. If the bilinguals in Experiment 1b were also activating the English translations of the Spanish words presented alongside the Dutch vocabulary during the study task, this convoluted route of translation should have produced much longer naming latencies for the Spanish words during the study task. Hence, these results suggest that bilinguals learning foreign language vocabulary words via L2 translations may be experiencing an advantage relative to monolinguals for different reasons than do bilinguals learning via L1 translations. This possibility will be returned to in the General Discussion.
CHAPTER 5: EXPERIMENT 2 – ARE BILINGUALS ADVANTAGED ON TASKS OF IMPLICIT LEARNING AND PATTERN RECOGNITION?

In light of the results of Experiments 1a and 1b, which demonstrated that bilinguals learning foreign language vocabulary via either L1 or L2 translations have an advantage relative to monolinguals, the underlying mechanism(s) supporting this advantage may be much more general than previous hypothesized. The goal of Experiment 2 was to further examine the possibility that bilinguals may be advantaged for multiple types of new learning, including implicit learning/processing. Though implicit learning is often characterized as the inverse of explicit learning, with implicit learning thought to be more akin to native language learning and explicit learning as what is done in second language learning, as an effortful and deliberate process (e.g., DeKeyser, 2003). However, the two frequently work in tandem, as even in explicit learning contexts, implicit learning can occur (Haider & Frensch, 2005). Hence, one possibility that can be tested is whether bilinguals are advantaged on tasks of implicit learning/processing, which would allow for the possibility that such an advantage underlies the bilingual advantage in foreign language vocabulary learning (though this would still need to be tested directly). To this end, in Experiment 2, a group of English-Spanish bilinguals—a group that has previously demonstrated a bilingual advantage in foreign language vocabulary learning—and monolinguals were tested on a series of implicit learning/processing tasks in both auditory (statistical learning) and visual (dot-motion discrimination) tasks.
5.1 Experiment 2

5.1.1 Method.

5.1.1.1 Participants. Twenty-one functionally monolingual English speakers and 14 native English speakers who had learned Spanish as a second language were recruited for Experiment 2. Data from four monolingual participants were discarded due to technical problems. The results presented for Experiment 2 reflect the remaining data from 17 monolingual and 14 bilingual participants. The English-Spanish bilinguals recruited for this study were predominantly late learners of Spanish (mean L2 onset, 13.6; range 8 to 19; thirteen of the fourteen total bilingual participants reporting, with only one of the 14 bilingual participants first exposed to Spanish earlier than age 12).

All participants were recruited from either the campus or the outlying areas surrounding the Pennsylvania State University, and were compensated $10 per hour for their participation.

5.1.1.2 Materials and Procedure. Experiment 2 recorded behavioral data only, and consisted of a single session that typically lasted between an hour a half to two hours. Participants first listened to a 10-minute language stream of nonsense syllables (described below), followed by a 32-item two-alternative forced choice test. They then completed the Operation-Span task, the AX-continuous performance task (Rosvold, Mirsky, Sarason, Bransome Jr., & Beck, 1956) and the Boston Naming test in English (Kaplan, Goodglass, & Weintraub, 2001). After the Boston Naming test, participants listened to a second 10-minute language stream of nonsense
syllables (again, described below), followed by another 32-item two-alternative forced choice test. Finally, participants completed a dot-motion discrimination task (Green, Pouget, & Bavelier, 2010). Bilingual participants also completed a Spanish version of the Boston Naming test at the end of the experimental session.

5.1.1.2.1 Statistical Learning Task. The materials for both presentations of the statistical learning paradigm were adapted from Gebhart, Aslin, and Newport (2009) and Newport and Aslin (2004). Hence, all synthesizing and acoustic properties of the languages used in the statistical learning task in Experiment 2 were identical to those reported. After pilot testing of the two language streams utilized by Gebhart et al. (2009) determined that the two languages were not equally learnable after 5 minutes of isolated exposure (i.e., with no exposure to additional language streams; learnability of the two languages was 90.6% and 65%, respectively), a language stream created by Newport and Aslin (2004) was also used in additional pilot testing, and determined to be a better match in learnability for one of the language streams used by Gebhart et al. (2009), with test performance at 75.5%. Hence, one of the languages constructed 16 trisyllabic words by using two 3-consonant frames that inserted one of two vowels in each of the 3 positions (see Table 5.1). Conversely, the other language consisted of 16 trisyllabic words by using two 3-vowel frames that inserted one of two consonants in each of the 3 positions (again, see Table 5.1).
Table 5.1. Design of the two counterbalanced languages in Experiment 2.

<table>
<thead>
<tr>
<th>Frames</th>
<th>Slot 1 Fillers</th>
<th>Slot 2 Fillers</th>
<th>Slot 3 Fillers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>_a_u_e</td>
<td>[d_]</td>
<td>[b_]</td>
</tr>
<tr>
<td></td>
<td>_o_i_a_e</td>
<td>[p_]</td>
<td>[t_]</td>
</tr>
<tr>
<td>Language 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t_d_k_e</td>
<td>[_a_e]</td>
<td>[_i_u]</td>
</tr>
<tr>
<td></td>
<td>b_p_g</td>
<td>[_e]</td>
<td>[_o]</td>
</tr>
</tbody>
</table>

With this construction, both languages preserved the following dependencies:

- Transitional probabilities between vowels in the vowel-frame language (Language 1) were 1.0 within words, and 0.5 between words, and transitional probabilities across consonants in the consonant-frame language (Language 2) were also 1.0 within words and 0.5 across words. Additionally, for both languages, transitional probabilities of syllables both within and across words was 0.5. Words within each language were then pseudorandomized and repeated across a five minute language stream, such that words never appear consecutively within the stream.

Finally, two 10 minute streams were constructed: one with the 5 minutes of the vowel-frame language stream followed by 5 minutes of the consonant-frame language stream, and the other with the reverse. No pauses or other explicit cues separated the two 5-minute streams.

For approximately half of the participants, the vowel-frame language served as Language A, or the first language heard during the first familiarization and the second language heard during the second familiarization, and for the other half of the participants the consonant-frame language served as Language A. Likewise, for
those participants who heard the vowel-frame language as Language A, they then heard the consonant-frame language as Language B, and vice versa.

The schematic for the presentation of both presentations of the statistical learning task is presented in Figure 5.1.

![Figure 5.1](image)

Figure 5.1. The statistical learning task procedure, adapted from Gebhart et al. (2009).

Participants listened to two consecutive and uninterrupted 5-minute exposures of the two aforementioned language streams, for a total of 10 minutes of familiarization. First, the participants would hear a 5-minute speech stream consisting of all 16 possible words from one of the languages previously described (i.e., either the vowel-frame or consonant-frame language) as Language A. Following the initial 5-minute familiarization, participants then heard an additional 5-minute familiarization consisting of all 16 possible words from the other of the two artificial languages constructed that they had not yet been exposed to; hence, all participants heard 5 minutes of one language followed by 5 minutes of the other language. No discernible cues were present to alert participants to the change in language inputs.

Immediately following the 10 minutes of familiarization (i.e., Language A then B, or Language B then A), the participants were given a two-alternative forced-choice test with one word and one part-word played per test item, and were asked to choose with a button press which item was more consistent with the sound
stream they had been listening to. The test consisted of 32 forced-choice items, with 16 items consisting of words and part-words derived from Language A, and 16 from Language B. Part-words also consisted of three syllables that had been present in the inventory, but contained two syllables from one of the 16 words, and one syllable from another.

Following the English Boston Naming test, a second statistical learning task was administered, consisting again of 10 minutes of familiarization followed by another 32 item two-alternative forced-choice test. The second familiarization was identical to the first, except the presentation order of the languages was switched: the first presentation consisted of Language A followed by Language B, and the second presentation consisted of Language B followed by Language A.

5.1.1.2 Operation-Span Task. Following this first familiarization and test tasks described above, the participants completed the Operation-Span task, which was identical to that used in Experiments 1a and 1b.

5.1.1.3 AX-CPT. After completing the Operation-Span task, a variation of the AX-CPT task was administered, adapted from Morales, Gómez-Ariza, and Bajo (in press), which was itself adapted from Ophir, Nass, Wagner, and Posner (2009). In this version of the AX-CPT task, every trial consisted of five letters presented, each for a duration of 300 milliseconds followed by an inter-trial interval of 1000 milliseconds on a black background in the following sequence: a cue (in red), three distracters (in white), and a probe (in red). The participants were asked to respond by pressing a "no" key on a button box for every probe and every distracter. The
response for the probe, however, was contingent both on the probe itself as well as the cue. Cue letters were randomly selected from all letters of the alphabet, save "X," "Y", and "K"; the former due to its identity as the target probe letter, and the latter two due to their visual similarity with the target probe letter. Similarly, the probe letters were randomly selected from all letters of the alphabet, save "A," "Y," and "K"; the former due to its identity as the target cue letter, and the latter two due, again, to their visual similarity with the target probe letter. The distracter letters were also randomly selected from all letters of the alphabet, excepting "A," "K," "X," and "Y" for the reasons aforementioned. The task consisted of 100 trials of four trial types presented: AX trials (70%), BX trials (10%), AY trials (10%), and BY trials (10%). Examples of these trial types can be seen in Figure 5.2.
Figure 5.2. The four trial types in the AX-CPT task: (A) An AX trial, (B) An AY trial, (C) A BX trial, and (D) a BY trial.

If the cue was the letter "A" and the probe was the letter "X", the participants were asked to respond with the "yes" on the button box (AX trials; (A) Figure 5.2). If the cue was the letter "A" but the probe as any letter other than "X", participants were instructed to respond to the probe with a "no" response (AY trials; (B) Figure 5.2). If the cue was any letter other than "A," and the probe was an "X," participants were still instructed to respond to the probe with a "no" response (BX trials; (C) Figure 5.2). Finally, if the cue was any letter other than "A" and the probe was any letter other than "X," participants were again instructed to respond to the probe with a "no" response (BY trials; (D) Figure 5.2).

5.1.1.2.4 Boston Naming Task. Upon completion of the AX-CPT task, participants were administered the Boston Naming test (Kaplan et al., 2001) as a measure of English proficiency and lexical access, which consisted of 60 black line drawings on a white background in descending order of frequency (i.e., high frequency pictures first, low frequency pictures last). Participants were asked to name each picture out loud into a microphone as quickly and accurately as possible. Pictures were presented on the screen for up to 5 s, or until their vocal response cued a voice trigger.

Bilingual participants were asked to complete the Boston Naming test again at the end of the experimental session, this time in Spanish, as a measure of proficiency and lexical access. All aforementioned parameters of the English Boston
Naming test were identical to that in the Spanish version, save for naming being done in Spanish. The results of the Spanish Boston Naming test will not be presented here, as the focus of Experiment 2 was on monolingual and bilingual comparisons.

5.1.2.5 Dot-Motion Discrimination Task. The final task administered to the monolingual participants (and the penultimate task for bilingual participants) was the dot-motion discrimination task (DMDT), adapted from that used by Green, Pouget, and Bavelier (2010). The task consisted of circular arrays of white dots in motion on a black background presented for a maximum duration of 2 seconds. On each trial, some percentage of the dots were moving a coherent direction (either right or left), which was varied by condition: 0.8%, 1.6%, 3.2%, 6.4%, 12.8%, 25.6%, and 51.2%. Fifty trials were presented to each participant per level of motion coherence, for a total of 350 experimental trials.

5.1.2 Results. The goal of Experiment 2 was to compare monolingual and bilingual performance on a series of implicit processing tasks—one linguistic, the other visual—and to determine whether the previously reported bilingual advantage in foreign language vocabulary learning for bilinguals may be an advantage in learning that would extend to less explicit learning contexts, such as statistical learning and dot-motion discrimination. Hence, the two groups' performance is compared across each task.
Participant descriptive statistics, including Operation-Span score, mean naming latencies and accuracy on the English version of the Boston Naming test, can be found in Table 5.2 below.

Table 5.2. Mean age, age of L2 onset, L1 self-ratings, L2 self-ratings, Operation-Span score, and mean naming latencies and accuracy on the Boston Naming test for participants in Experiment 2 (standard deviations in parentheses, significant differences between groups are denoted with *).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age</th>
<th>Age of L2 Onset</th>
<th>L1 Self-Rating</th>
<th>L2 Self-Rating*</th>
<th>O-Span Score*</th>
<th>BNT English RT</th>
<th>BNT English Accuracy*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolinguals</td>
<td>17</td>
<td>22.4 (5.3)</td>
<td>12.9 (4.0)</td>
<td>9.3 (0.8)</td>
<td>2.9 (2.1)</td>
<td>45.6 (8.4)</td>
<td>912 (142)</td>
<td>80.4% (10.4)</td>
</tr>
<tr>
<td>Bilinguals</td>
<td>14</td>
<td>24.6 (6.1)</td>
<td>13.6 (2.5)</td>
<td>9.6 (0.6)</td>
<td>7.0 (1.7)</td>
<td>54.5 (3.7)</td>
<td>925 (85)</td>
<td>88.5% (6.5)</td>
</tr>
</tbody>
</table>

Separate one-way ANOVAs conducted for each of the variables presented in Table 5.2 revealed no differences between groups for age, age of L2 onset, or L1 self-ratings (all ps > .10). Unsurprisingly, bilinguals rated their knowledge of their L2 higher than did monolinguals (F(1, 29) = 30.29, p < .001). Additionally, like the bilinguals reported in Experiments 1a and 1b, the bilinguals in Experiment 2 outperformed the monolinguals on the Operation-Span task (F(1, 29) = 13.46, p < .01). Finally, naming latencies for the Boston Naming test were not statistically significant between groups (F < 1), even though the Boston Naming test contains many very low-frequency exemplars (e.g., yoke, abacus) to be named, the bilinguals were no slower than the monolinguals in naming these pictures in their dominant language. The frequency-lag hypothesis (Gollan et al., 2011; Gollan, Montoya,
Fennema-Notestine, & Morris, 2005) would have predicted slower naming latencies for the bilinguals, who must split their time between languages, making exemplars in both languages functionally less frequent than for monolinguals. However, the bilinguals did accurately name more pictures than did the monolinguals ($F(1, 29) = 6.38, p < .05$). As this was not a specific prediction made between the groups, it will be returned to in the General Discussion.

**5.1.2.1 Statistical Learning Task.** First, a 2 (group: monolinguals, bilinguals) by 2 (learning phase: first, second) by 2 (language: A, B) by 2 (vowel-frame language first, consonant-frame language first) mixed factor ANOVA was conducted in order to rule out the possibility that the learnability of the two language streams used (i.e., vowel-frame and consonant-frame languages) was unequal. No main effect or interactions involving the counterbalancing factor approached significance (All $p$s > .10). Accuracy performance (in percent correct) on the two-alternative forced choice test given at the end of each statistical learning phase was submitted to a 2 (group: monolinguals, bilinguals) by 2 (learning phase: first, second) by 2 (language: A, B) mixed factor ANOVA. Recall that during the first familiarization phase, all participants heard 5 minutes of exposure to Language A followed by 5 minutes of Language B. The second familiarization phase then switched the order, presented 5 minutes of Language B followed by 5 minutes of Language A. The accuracy performance for both groups across both languages and both familiarization presentations can be seen in Figure 5.3.
Figure 5.3. Percentage of test items correctly identified for both statistical learning tasks in Experiment 2 for bilinguals and monolinguals.

The ANOVA revealed a marginally significant main effect of language ($F(1, 29) = 3.62, p = .07, \eta^2 = .11$), such that language A was better learned by all participants regardless of learning phase. The interaction between language and learning phase approached significance ($F(1, 29) = 2.13, p = .16, \eta^2 = .07$), as, if anything, Language A was best learned after the first familiarization, and performance on Language A test items was slightly less accurate after the second familiarization, while performance for Language B items was approximately equal at both tests. No other main effects or interactions approached significance (all $Fs < 1$).

Attending to the results of the first familiarization phase only for the moment, the primacy effect reported by Gebhart et al. (2009) for languages presented without an explicit cue to the language switch was replicated for both
monolingual and bilingual participants. However, what is most striking about these results is that the language first heard during the first familiarization of learning was still the better-learned language at the second familiarization test, even though it was the second language heard in this second phase. Independent of language experience (i.e., being bilingual or monolingual), it is difficult to learn a second language (i.e., language B) in a first familiarization stream if there is no explicit boundary cue between the languages. Moreover, and also independent of language experience, this learning of language A persists for at least an hour (the amount of time that elapsed between the first and second familiarization phases). Critically, however, monolinguals and bilinguals did not differ in their ability to learn either language at either learning phase.

**5.1.2.2 AX-CPT.** Following the procedure outlined by Morales et al. (in press), all correct trials that were either faster than 100ms or slower than 1000ms were removed from reaction time analyses. Reaction time data for both groups in all four conditions can be found in Figure 5.4.
The remaining reaction time data for probe responses (i.e., the final letter in the 5 letter sequence) were submitted to a 2 (group: monolingual, bilingual) by 4 (condition: AX, AY, BX, BY) mixed factor ANOVA, which revealed a main effect of condition \( (F(3, 87) = 77.86, p < .001, \eta_p^2 = .73) \). Planned Bonferroni-corrected pairwise comparisons revealed that all four conditions significantly differed from one another (all \( p_s < .05 \)). BX trials received the fastest responses, followed by BY trials, followed by AX trials, followed by AY trials. No main effect of group was observed \( (F < 1) \), though the group by condition interaction approached significance \( (F(3, 87) = 2.01, p = .14, \eta_p^2 = .07) \). This interaction seems to have been driven by the fact that, although the two groups were similarly fast in their performance for each condition, monolinguals were slightly slower on the BY trials than were bilinguals.
Accuracy data for both groups in all four conditions can be found in Figure 5.5.

Figure 5.5. AX-CPT accuracy data for monolinguals and bilinguals.

The accuracy data for probe responses on the AX-CPT were submitted to an additional 2 (group: monolingual, bilingual) by 4 (condition: AX, AY, BX, BY) mixed factor ANOVA, which revealed a main effect of condition \( F(3, 87) = 4.91, p < .01, \eta^2_p = .15 \). Planned Bonferroni-corrected pairwise comparisons revealed that performance in the BY condition was better than in the AY condition \( t(1) = 1.84, p < .05 \). Additionally, performance in the AX condition was slightly better than performance in both the AY and BX conditions, though both effects were marginally significant \( t(1) = 2.76, p = .06; t(1) = 2.76, p = .06 \), respectively). Finally, the main effect of group was significant \( F(1, 29) = 4.23, p < .05, \eta^2_p = .13 \), as overall, bilinguals
outperformed monolinguals across all conditions of the task. The group by condition interaction did not approach statistical significance.

Unlike Morales et al. (in press), who found that bilinguals were slower to perform the AX-CPT than monolinguals, no differences in speed of performance were found between monolinguals and bilinguals in Experiment 2, although, if anything, the trend in the reaction time data was toward bilinguals being slightly faster than monolinguals. Although the main effect of group did not approach statistical significance, the group by condition interaction did, as bilinguals were specifically faster than monolinguals in the BY condition. Additionally, while Morales et al. found a bilingual advantage for accuracy in the AY condition only, the bilinguals in Experiment 2 outperformed the monolinguals overall, not just in the AY condition (although numerically, the greatest difference between the groups existed for AY trials). Hence, these data can reasonably be interpreted as support for the notion that the late bilinguals tested in Experiment 2 are advantaged at both proactive (BX relative to BY trials) and reactive (AY relative to BY trials) inhibitory control.

The data from the bilinguals tested in Experiment 2 have thus far demonstrated an interesting set of results that support the idea that both bilinguals and bilingual advantages are not only not binary categorizations (e.g., Luk & Bialystok, 2013), but vary across multiple dimensions according to numerous factors, such as frequency of usage, proficiency, age of acquisition, etc. For instance, the bilinguals in Experiment 2 exhibited a larger working memory span relative to
the monolinguals (as evidenced by higher Operation Span scores), and enhanced inhibitory control relative to the monolinguals as evidenced by the AX-CPT. However, this same group that demonstrated these advantages did not exhibit any discernible advantages on the statistical learning task. In order to further assess what set of advantages this bilingual group may have had, performance on the dot-motion discrimination task is investigated below.

5.1.2.3 Dot-Motion Discrimination Task. Following Green et al. (2010), the data from the dot-motion discrimination task were submitted to two mixed-factor ANOVAs: one for the accuracy data, and other for the reaction time data. The accuracy and reaction time data can be found in Figures 5.6a and 5.6b, respectively.

Figure 5.6a. Accuracy rates across all seven levels of dot-motion coherence for monolinguals and bilinguals in Experiment 2.
Figure 5.6b. Mean reaction times across all seven levels of dot-motion coherence for monolinguals and bilinguals in Experiment 2.

The first 2 (group: bilingual, monolingual) by 7 (coherence levels: 0.8%, 1.6%, 3.2%, 6.4%, 12.8%, 35.6%, and 51.7%) mixed-factor ANOVA revealed a main effect of dot coherence level ($F_{6, 174} = 210.99, p < .001$), as performance at low levels of coherence were much less accurate than performance at high levels of coherence. Planned Bonferroni-corrected pairwise comparisons revealed significant differences between all levels of coherence except between the lowest two levels (0.8%, 1.6%) and between the highest two levels (35.6%, 51.7%; all other $p$s < .01). Similarly to the Green et al. study with action video-game players and non-players, no main effect of group or group by coherence level interaction emerged for the accuracy data (both $F$s < 1).
For the reaction time analyses of the dot-motion discrimination task, trials faster than 100ms and slower than 1900ms were excluded from reaction time analyses. These data were also submitted to a 2 (group: bilingual, monolingual) by 7 (coherence levels: 0.8%, 1.6%, 3.2%, 6.4%, 12.8%, 35.6%, and 51.7%) mixed-factor ANOVA, which again revealed a main effect of dot-motion coherence ($F(6, 174) = 88.49, p < .001, \eta_p^2 = .75$), as performance was slowest for the lowest levels of dot-motion coherence, and fastest for the highest levels. Planned Bonferroni-corrected pairwise comparisons revealed that all levels of dot-motion coherence differed from one another (becoming faster at increasing levels of dot-motion coherence), except among the first three levels (0.8%, 1.6%, and 3.2%), where performance was similarly slow. While numerically the bilinguals were faster than the monolinguals overall, the main effect of group did not reach statistical significance ($F(1, 29) = 1.77, p = .19, \eta_p^2 = .06$), and neither did the group by coherence level interaction ($F < 1$). However, power in this task was quite low (observed power = .25). Although the total number of subjects per group was similar to that of Green et al. (11 action-video game players and 12 non-players), the total number of observations per condition was quite different. In the study by Green et al., not only were the number of data points doubled (i.e., 100 trials per dot-motion coherence condition, rather than 50), their participants completed 3 practice runs across a 3-day time span, including 100 trials for each of the 7 coherence conditions at each practice run on each of the three days. The reported data by Green et al. were collected from a fourth run on a fourth day.
Such a time-intensive version of this task was not possible within the context of Experiment 2, but the results reported here at least suggest the possibility of a bilingual advantage on a dot-motion discrimination task relative to monolinguals. However, the results of Experiment 2 do not suggest a perfect parallel to the results reported by Green et al. First, while the non-video game players performed the task with similar speeds to that of the participants reported in Experiment 2, the action video game players were a bit faster than either group, especially for the lower levels of dot-motion coherence. Hence, while both bilinguals and action video game players may have benefits in attentional processing and/or probabilistic inference, the data suggest that these benefits are not exactly the same for the two groups. With more training over multiple days, it remains possible that the performance reported here from the bilingual and monolingual groups would change. However, a somewhat more likely possibility is that both bilinguals and action video game players may have benefits to their executive function networks, but that these networks are being exercised very differently. This notion of various types of experience tuning executive and linguistic networks differentially (coined the adaptive control hypothesis by Green & Abutalebi, 2013) will be returned to in the General Discussion.
CHAPTER 6: GENERAL DISCUSSION

The goal of the present dissertation was to further investigate the nature of the bilingual advantage in foreign language vocabulary learning by considering the possibility that bilinguals may have an advantage in learning more generally. Additionally, this advantage in learning may be related to other reported advantages in executive function, or it may be an advantage that is independent of documented effects of bilingualism on executive function. In order to experimentally test these ideas, bilinguals and monolinguals were compared over a variety of language, executive function, and learning/implicit processing tasks.

6.1. Summary of Findings

6.1.1. Summary of Experiment 1a. The goal of Experiment 1a was to first replicate the bilingual advantage in foreign language vocabulary learning via L1 translations, and then to examine whether neurophysiological evidence (as indexed by ERPs) provided any additional evidence regarding the time course and/or development of the advantage. The results of Experiment 1a did produce evidence for a bilingual advantage in foreign language vocabulary learning for bilinguals learning via L1 translations. Both monolinguals and bilinguals demonstrated evidence of learning in the behavioral record, but the bilingual group showed the strongest evidence of learning. Additionally, the bilinguals in Experiment 1a also exhibited longer naming latencies during the study task, as did the bilingual L1 learners test by Bogulski and Kroll (under review), though the magnitude of the effect in Experiment 1a was smaller than that reported by Bogulski and Kroll. These
longer naming latencies are consistent with the inhibitory hypothesis they propose, which argued that the bilingual advantage in foreign language vocabulary learning depends on experiencing interference from the dominant language during study that must be inhibited. The bilinguals tested in Experiment 1a also demonstrated a marginally significant advantage in working memory span relative to the monolinguals. Hence, working memory may be related to advantages in foreign language vocabulary learning. Additionally, the ERP evidence revealed that the monolinguals showed very little evidence (if any) of differential sensitivity to Dutch-like nonwords relative to real Dutch words as indexed by ERPs (see Figure 4.5). In contrast, the bilinguals did show a slight N400 to the Dutch-like nonwords relative to the real Dutch words (again, see Figure 4.5).

6.1.2. Summary of Experiment 1b. Experiment 1b aimed to further investigate the role of the context of language learning on the bilingual advantage in foreign language vocabulary. Specifically, Bogulski and Kroll (under review) had argued previously that the bilingual advantage in foreign language vocabulary learning depended on learning via dominant language translations. They further suggested that the presence of the dominant language context in learning activated inhibitory networks that led to slower naming latencies during the study task, but better learning in the test task. In order to more directly test this hypothesis, another group of English-Spanish bilinguals were taught and tested on a group of Dutch words using behavioral and ERP measures. However, this group of bilinguals—unlike those reported in Experiment 1a—learned foreign language
vocabulary via L2 (Spanish) translations. This group was predicted not to demonstrate the bilingual advantage in foreign language vocabulary learning, as the language of translations was no longer the dominant one. However, the results of Experiment 1b did not confirm this prediction. These bilingual L2 learners not only exhibited a bilingual advantage in foreign language vocabulary learning relative to the monolingual learners tested in Experiment 1a, but their naming latencies during the study task were comparatively fast relative to those of the bilingual L1 learners tested in Experiment 1a. If the bilingual L2 learners had been engaging in similar processing as the bilingual L1 learners, then they might have been expected to demonstrate even slower naming latencies during the study task relative to the bilingual L1 learners, as they were also naming in their non-dominant language. All the more surprising were the ERP results: the bilingual L2 learners showed even larger N400 effects than did the bilingual L1 learners, and this effect was most like those reported in typical N400 elicitations (i.e., centro-parietal distribution).

The results of Experiment 1b failed to support the inhibitory hypothesis, as bilinguals learning via L2—a language they have little experience inhibiting. These results supported a much more complex interaction between language experience and context of language learning, which will be returned to in the final discussion.

6.1.3. **Summary of Experiment 2.** The aim of Experiment 2 was to test another group of bilinguals and monolinguals and compare their performance on a number of executive function and learning/implicit processing tasks. In making these comparisons, the goal of Experiment 2 was to address the possibility that
bilinguals are advantaged at learning more generally, and it is this general learning advantage that drives an advantage in foreign language vocabulary learning. The statistical learning task could be characterized as a linguistic implicit learning task in which attentional (or inhibitory) control may have played a role, given the presence of multiple language streams. However, performance between the bilingual and monolingual groups was statistically indistinguishable. The dot-motion discrimination task had previously been used by Green et al. (2010), where an advantage was found for action video game players relative to non-players. Though bilinguals and video game players are clearly differentially skilled, it has been argued that their respective expertise makes extensive—but differential—use of an executive function network. Although the results are somewhat speculative, given the smaller number of observations per subject, per condition, bilinguals trended toward faster performance on the task relative to monolinguals. However, Green et al. reported the largest differences between video game players and non-players for the lowest levels of dot-motion coherence, while the bilinguals in Experiment 2 seemed faster than the monolinguals overall. This overall speed of processing advantage is reminiscent of Hilchey and Klein’s (2011) recent review of bilingual advantages reported on Simon, Stroop, and Flanker tasks. Hilchey and Klein argue that the advantage bilinguals have relative to monolinguals in not inhibitory control, but in overall speed of processing, which was the trend in the dot-motion discrimination task data.
Additionally, bilinguals outperformed the monolinguals on an AX-CPT. Specifically, bilinguals were faster than monolinguals to respond in the most demanding conditions. However, unlike Morales et al. (in press), who reported a bilingual advantage exclusively for trials requiring reactive inhibition, the bilinguals in Experiment 2 had advantages for both the types of the trials that required reactive as well as proactive inhibition. Furthermore, these bilinguals also exhibited an advantage in working memory and higher accuracy on a difficult picture-naming task. Hence, clearly the bilinguals tested in Experiment 2 had advantages relative to the monolinguals, but these advantages did not extend to enhanced statistical learning or even implicit processing (as evidence by the dot-motion discrimination task), where the advantage trended toward an overall processing speed advantage.

What is particularly noteworthy about these results is that between the two implicit processing tasks—statistical learning and dot-motion discrimination—the bilinguals tested in Experiment 2 were, if anything, slightly better than the monolinguals at the visual, non-linguistic task, not the auditory, linguistic one. It was not the case that these particular bilinguals performed exactly like monolinguals; on the contrary, they outperformed the monolinguals on the AX-CPT. Moreover, they also exhibited larger a larger vocabulary (as indexed by the Boston Naming test).

6.2 Conclusions. Across three experiments, evidence for bilingual advantages were observed, but not for all tasks. While some advantages found clear support, such as foreign language vocabulary learning and AX-CPT, others were either absent or marginally significant, such as working memory and dot-motion
discrimination. Moreover, the central target of the investigation of bilingual advantages—foreign language vocabulary learning—was observed with approximately equal levels of success for both for bilinguals learning via L1 translations and for bilinguals learning via L2 translations. However, only the bilingual L1 learners exhibited longer naming latencies for the presented translations during the study task. Moreover, the ERP results revealed the greatest sensitivity to Dutch-like nonwords relative to real Dutch words for the bilingual L2 learners, though both groups of bilingual learners showed greater sensitivity than did the monolinguals in the ERP record, as evidenced by the N400 component to the nonwords relative to the words (see Figure 4.5).

In Experiment 2, the bilinguals did not outperform the monolinguals on the statistical learning task, although the primacy effect reported by Gebhart et al. (2009) was replicated. Hence, neither group was insensitive to the manipulation of presenting separate language streams, but both were equally likely to best learn the statistics of the first language presented and retain that learning throughout the experiment through the second language stream familiarization, in which the first language heard in the first familiarization was now the second language heard without a cue to distinguish the shift. This group of bilinguals was drawn from the same population (and approximately half of them were literally the same participants as in Experiment 1b) was the same group that exhibited a bilingual advantage in foreign language vocabulary learning.
These results further support the idea that bilingualism is not a categorical variable, but rather varies on a multi-dimensional continuum (Luk & Bialystok, 2013). What is striking about these findings is how even among the same types of bilinguals, the context of the task environment can dramatically alter the results. For example, changing the language of translation in the study task from English to Spanish (as in Experiments 1a and 1b) changed the outcome, with both bilingual groups exhibiting an advantage, but the bilingual L2 learners showing faster naming latencies at study and greater neurophysiological sensitivity to the nonwords relative to the words.

Given that a bilingual advantage in foreign language vocabulary learning has been reported for both early (Kaushanskaya & Marian, 2009a, 2009b) and late bilinguals (Bogulski & Kroll, under review; Van Hell & Mahn, 1997), as well as for bilinguals learning via L1 and L2 translations, the foreign language vocabulary learning advantage seems to be dependent on bilingual language usage, as these quite disparate groups exhibiting this advantage all have in common the fact that they know and use multiple languages. However, Bogulski and Kroll (under review) failed to find evidence that bilinguals learning via L2 translations exhibited this advantage. The bilingual L2 learners tested by Bogulski and Kroll and those tested in Experiment 1b differed in one critical aspect: those tested by Bogulski and Kroll were immersed in their L2 environment, while those tested in Experiment 1b were immersed in their native language environment. Hence, while knowledge of multiple languages may be critical for exhibiting an advantage in foreign language vocabulary
learning, it may also be important for the language learning context to take place in an L1 environment. As previously reviewed, effects of L2 immersion on L1 processing have been found at both the lexical (e.g., Linck et al., 2009) and grammatical (Dussias & Sagarra, 2007) levels. However, the impact of L1 or L2 immersion on foreign language vocabulary learning in bilinguals has not yet been systematically investigated. The results of the present set of experiments suggest that L1 immersion in the environment (outside of the laboratory) may differentially engage cognitive networks that are further differentially engaged depending on the language through which the new vocabulary is being learned. Whether this is also the case for L2 immersion (i.e., differential success in foreign language vocabulary learning relative to monolingual performance) depending on foreign language vocabulary being acquired via L1 or L2 translations remains a topic for future investigation.

Experiments 1a and 1b not only compared bilingual and monolingual performance on a vocabulary learning task behaviorally, but ERPs were collected in order to examine the possibility that the bilingual advantage in foreign language vocabulary learning extends to a neurophysiological sensitivity. The results reported in these experiments suggest that such sensitivity exists for bilinguals. However, this sensitivity is not a perfect index of learning that is taking place behaviorally. Unlike McLaughlin et al. (2004), no group tested in the present set of experiments showed ERP sensitivity during learning of the absence of behavioral evidence. This is likely due to the laboratory nature of the study environment, which
differed from that used by McLaughlin et al. in a necessarily abbreviated vocabulary list and testing phases that were much closer in time. However, some interesting effects in the ERP record did emerge. For example, while both bilingual L1 and bilingual L2 learners exhibited similar advantages at the behavioral test, their ERP signatures differed, as the bilingual L2 learners showed enhanced sensitivity to Dutch-like nonwords relative to the Dutch words. Whether or not this sensitivity would eventually lead to differential behavior is an interesting question that cannot be answered with the current data set. However, if such differences did ultimately emerge at some point during learning, this would have profound consequences for language learning in the real world, as well as for neurophysiological models of language processing.

At present the evidence suggests that both language usage (i.e., bilingualism) and context of learning (i.e., via L2 translations) can impact how the brain is processing new language input. However, ERP signatures do not appear to be concomitant with foreign language vocabulary learning, as the monolingual learners clearly learned some of the foreign language vocabulary while showing very little differential sensitivity to the nonwords and words in the ERP record. Part of the issue in the present methodological design was that nonword trials were repeated, perhaps making the nonwords appear more "word-like" due to their increasing familiarity. Hence, more research is needed to determine how ERP sensitivity can or cannot predict vocabulary learning success for groups with different language experience. The results presented in Experiment 1a suggest that there may be
multiple routes available to learners, as both bilingual L1 learners and monolinguals learned some of the vocabulary, but with differing levels of success. The monolingual learners were likely engaged in differential processing than were the bilingual L1 learners. One possibility is that the demands of the study task could be changed to simulate a "bilingual" learning experience in monolinguals in order to attain greater success. If such a manipulation were successful, perhaps the cognitive networks of the monolingual learners would be differentially engaged, such that they would exhibit not only enhanced learning as indexed by behavioral measures, but also differential sensitivity in the ERP record for the nonwords relative to the words. Again, future research is necessary in order to ascertain the predictive relationship between ERP evidence and behavioral learning within the context of bilingual/monolingual comparisons in foreign language vocabulary learning.

Bilingual performance relative to monolinguals on the tasks of inhibitory control were somewhat mixed. Experiment 1a found partial overlap with those advantages reported by Luk (2008) on the flanker task, but Experiment 1b failed to find any bilingual advantages on the flanker task. However, the bilinguals in Experiment 2 exhibited a robust advantage on an AX-CPT, but these results differed from those reported for the early bilinguals tested by Morales et al. (in press). Hence, given that even among the relatively homogenous samples tested across the three experiments reported here, these effects were inconsistently identified. This suggests that the bilingual advantage in at least some aspects of executive function—such as inhibitory control, cognitive flexibility—are much more
susceptible to variations in language experience. Unlike the bilingual advantage in foreign language vocabulary learning, which has been reported for several types of bilinguals across multiple languages, advantages on the flanker task, AX-CPT, and even the Operation-Span task are found repeatedly, but inconsistently. This suggests that the profiles of the particular bilinguals tested must be tightly monitored in order to accurately assess what types of bilingual experience lead to enhanced executive function abilities.

Statistical learning is also a type of task that has found inconsistent support for bilingual advantages (Bartolotti et al., 2011; Yim & Rudoy, 2013). Although statistical learning can be characterized as a language task likely recruiting language processing mechanisms, bilinguals (relative to monolinguals) do not robustly demonstrate an enhanced ability to learn statistical patterns, as they do to learn foreign language vocabulary. While both the bilinguals and monolinguals tested in Experiment 2 were able to learn some of the language input presented to them in a statistical learning task, they did not do so with differential success. This suggests that the two skills—statistical learning and foreign language vocabulary learning—relate at least partially on independent underlying mechanisms. Again, the mixed set of results found for bilingual advantages learning statistics, like the results found for bilingual advantages on AX-CPT and flanker tasks, may rely on particular mechanism(s) or networks that are differentially tuned depending on language experience across the lifespan.
Like the perceptual wedge hypothesis (Petitto et al., 2012) and the adaptive control hypothesis (Green & Abutalebi, 2013) propose, early bilingual experience is likely to differentially alter how linguistic and cognitive control networks are tuned, so as to optimally adapt to the particular input of the environment. These consequences are almost certainly different from those that occur for late bilinguals. However, the presence of several types of bilingual advantages for late bilinguals suggests that these systems are incredibly plastic. What is perhaps most interesting about the adaptive control hypothesis is that both late and early bilinguals sometimes show similar benefits (e.g., inhibitory control, such as the flanker task and AX-CPT). This suggests that early and late bilinguals are not categorically different, but rather have networks particularly tuned for their language experience and usage. Some of these adaptations must be similar for the two groups, as both groups must enhanced whatever support networks are necessary to facilitate the fluent use of two languages. These similarities lead to advantages such as that reported for foreign language vocabulary learning, which is reported for multiple types of bilinguals with very different language experience profiles.

Similarly, the action video game players who demonstrate an advantage on the dot-motion discrimination task relative to non-players are a set of individuals who have in common their propensity to play a type of game that demands success in attentional control and, extrapolating from the label “probabilistic inference,” statistical learning. The fact that the performance from the action video game players and the bilinguals tested in Experiment 2 is so similar on the dot-motion
discrimination task (i.e., both groups exhibit an overall speed advantage) at first seems surprising. However, both groups developed a type of cognitively demanding expertise later in life. And though differences exist between the two groups on the task (i.e., action video game players are especially fast to respond at the lowest levels of dot-motion coherence), this is likely due to the differential demands of their type of expertise relative to that of bilinguals. Perhaps the reason that Hilchey and Klein (2011) find an overall speed of processing advantage for bilinguals in the flanker, Simon, and Stroop tasks is because speed of processing is a bilingual advantage much more like that found for foreign language vocabulary learning: multiple types of bilinguals are advantaged because skilled performance relies on the same mechanisms that support knowing and using multiple languages. Other effects, such as an advantage in inhibitory control, may rely on more specific aspects of bilingualism, which may be lost in a meta-analysis in collapsing across a variety of bilingual profiles.

The current set of experiments, perhaps for the reasons outlined above, found inconsistent support for a bilingual advantage in executive function. While it is certainly worthwhile to further examine the relationship between performance on all of these variables (such as flanker performance and foreign language vocabulary learning; AX-CPT and statistical learning), it is beyond the scope of the present dissertation to examine these in detail. With such analyses, and likely with many additional participants, the precise relationships between these skills may be more adequately investigated. For now, it appears that the bilingual advantage in
foreign language vocabulary learning is at least partially separable from that of advantages reported for executive function, as the bilingual advantage in foreign language vocabulary learning has been found for many different types of bilinguals, and, particular to this set of experiments, influenced by the language of translation. Further research is necessary in order to more specifically determine how factors such as L1 or L2 immersion differentially engage the networks responsible for foreign language vocabulary learning, and possibly other types of learning as well.

The account that best fits the results presented here is one proposed by Green and Abutalebi (2013). Their adaptive control hypothesis emphasizes the role that different types of language experience can have on domain-general control processes and neural networks. This account not only allows for different types of bilinguals to exhibit different advantages, but also allows for variation at the level of the individual speaker to impact these networks. This experiential tuning may even have a physical manifestation, such as in the anterior cingulate cortex, which is a structure involved in cognitive control (Abutalebi et al., 2012). Consider the results of Experiments 1a and 1b. Both groups of bilingual participants were drawn from the same population of native English speakers who had learned Spanish to moderate to high levels of proficiency later in life. However, manipulating the language through which they acquired foreign language vocabulary revealed different results. Such evidence might be interpreted as differential engagement of similarly tuned control networks as learning via L1 translations had dramatic consequences for study naming latencies, and learning via L2 translations had
similarly dramatic effects on neurophysiological sensitivity to the newly acquired vocabulary. While enhanced learning occurred for both groups relative to a monolingual group, their learning profiles as indexed by both ERP and behavioral evidence looked quite different. This suggests that the bilingual advantage in foreign language vocabulary learning is not an all-or-nothing prospect, but rather a manifestation of a finely-tuned network that can be altered dramatically either by changing aspects of the types of learners involved or by manipulating the context of learning.

This hypothesis also elegantly addresses somewhat tricky questions posed by research teams using the same tasks, testing slightly different groups of bilinguals, and reporting different results. For example, the late bilinguals in Experiment 2 showed a different pattern of performance on AX-CPT than did the bilinguals reported by Morales et al. (in press). These seemingly conflicting results suggest that early, balanced bilinguals may have differentially tuned their executive network relative to late, unbalanced bilinguals in order to support the language processing goals of being an early, balanced bilingual. Likewise, being a late, unbalanced bilingual imposes different cognitive demands that will also exercise the same cognitive network, but in a slightly different way. Not all types of cognitive experts are skilled at all cognitive tasks, and similarly, not all types of bilinguals have the same types of language or cognitive advantages. But rather, bilinguals provide a common population of experts with interesting variability that has differential cognitive and neural consequences.
Allowing for such highly complex interactions re-directs the question of "is the bilingual advantage in foreign language vocabulary learning just another manifestation of an advantage in executive function?" to "what individual and contextual variables predict better performance?" or "how can different types of language experience, or even bilingual experience, differentially tune cognitive and neural networks?". The short answer to these questions is that we do not yet know the answers. However, the evidence shown across Experiments 1a, 1b, and 2—in conjunction with previous research—suggests that seemingly small differences (i.e., learning via L1 or L2 translations) have large impacts, and that sometimes groups with very different life experience exhibit similar advantages (i.e., bilinguals and action video game players). Uncovering how best to exploit these networks may ultimately prove the most productive and important area of research in the years to come.
REFERENCES


doi:10.1080/713755891


doi:10.1017/S1366728906002847


# APPENDIX

Dutch words with their English and Spanish translations used in Experiment 1a and 1b

<table>
<thead>
<tr>
<th>Dutch word</th>
<th>English translation</th>
<th>Spanish translation</th>
<th>Dutch log frequency*</th>
<th>English log frequency*</th>
<th>Spanish log frequency†</th>
<th>Dutch Length</th>
<th>English Length</th>
<th>Spanish Length</th>
<th>Dutch-like Nonword</th>
<th>English/Spanish Cognate Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>aardig</td>
<td>nice</td>
<td>simpático</td>
<td>1.85</td>
<td>2.15</td>
<td>1.51</td>
<td>6</td>
<td>4</td>
<td>9</td>
<td>oosdig</td>
<td>possible cognate</td>
</tr>
<tr>
<td>adem</td>
<td>breath</td>
<td>respiración</td>
<td>1.94</td>
<td>1.74</td>
<td>1.91</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>arev</td>
<td>possible cognate</td>
</tr>
<tr>
<td>ambt</td>
<td>office</td>
<td>oficina</td>
<td>1.04</td>
<td>2.40</td>
<td>1.90</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>ibbt</td>
<td>cognate</td>
</tr>
<tr>
<td>arend</td>
<td>eagle</td>
<td>águila</td>
<td>0.48</td>
<td>0.85</td>
<td>1.46</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>urefd</td>
<td>noncognate</td>
</tr>
<tr>
<td>bedrijf</td>
<td>company</td>
<td>negocios</td>
<td>1.86</td>
<td>2.25</td>
<td>1.88</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>bevrilm</td>
<td>possible cognate</td>
</tr>
<tr>
<td>bestaat</td>
<td>exist</td>
<td>existe</td>
<td>2.03</td>
<td>1.20</td>
<td>2.44</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>beltaet</td>
<td>cognate</td>
</tr>
<tr>
<td>beurt</td>
<td>turn</td>
<td>turno</td>
<td>1.81</td>
<td>1.97</td>
<td>1.74</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>buird</td>
<td>cognate</td>
</tr>
<tr>
<td>bijeen</td>
<td>together</td>
<td>juntos</td>
<td>1.49</td>
<td>2.56</td>
<td>2.25</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>cojeon</td>
<td>noncognate</td>
</tr>
<tr>
<td>bodem</td>
<td>floor</td>
<td>suelo</td>
<td>1.67</td>
<td>2.21</td>
<td>2.64</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>toder</td>
<td>noncognate</td>
</tr>
<tr>
<td>boer</td>
<td>farmer</td>
<td>agricultor</td>
<td>1.66</td>
<td>1.49</td>
<td>0.60</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>pieg</td>
<td>noncognate</td>
</tr>
<tr>
<td>deugd</td>
<td>virtue</td>
<td>virtud</td>
<td>1.00</td>
<td>1.36</td>
<td>1.97</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>roulg</td>
<td>cognate</td>
</tr>
<tr>
<td>dienst</td>
<td>service</td>
<td>servicio</td>
<td>2.10</td>
<td>2.25</td>
<td>2.29</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>bianst</td>
<td>cognate</td>
</tr>
<tr>
<td>doos</td>
<td>box</td>
<td>caja</td>
<td>1.34</td>
<td>1.59</td>
<td>2.08</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>gios</td>
<td>noncognate</td>
</tr>
<tr>
<td>dorp</td>
<td>village</td>
<td>pueblo</td>
<td>1.99</td>
<td>2.13</td>
<td>2.72</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>ferp</td>
<td>possible cognate</td>
</tr>
<tr>
<td>draad</td>
<td>wire</td>
<td>hilo</td>
<td>1.28</td>
<td>1.54</td>
<td>1.90</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>drous</td>
<td>noncognate</td>
</tr>
<tr>
<td>druif</td>
<td>grape</td>
<td>uva</td>
<td>0.90</td>
<td>0.30</td>
<td>1.04</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>truib</td>
<td>noncognate</td>
</tr>
<tr>
<td>eed</td>
<td>oath</td>
<td>juramento</td>
<td>0.90</td>
<td>0.70</td>
<td>0.85</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>euw</td>
<td>noncognate</td>
</tr>
<tr>
<td>eend</td>
<td>duck</td>
<td>pato</td>
<td>1.08</td>
<td>0.60</td>
<td>1.00</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>eols</td>
<td>noncognate</td>
</tr>
<tr>
<td>emmer</td>
<td>bucket</td>
<td>cubo</td>
<td>1.11</td>
<td>1.11</td>
<td>1.41</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>imbek</td>
<td>possible cognate</td>
</tr>
<tr>
<td>ernaar</td>
<td>forward</td>
<td>adelante</td>
<td>1.45</td>
<td>2.08</td>
<td>2.32</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>alnaar</td>
<td>noncognate</td>
</tr>
<tr>
<td>Dutch word</td>
<td>English translation</td>
<td>Spanish translation</td>
<td>Dutch log frequency*</td>
<td>English log frequency*</td>
<td>Spanish log frequency†</td>
<td>Dutch Length</td>
<td>English Length</td>
<td>Spanish Length</td>
<td>Dutch-like Nonword</td>
<td>English/Spanish Cognate Status</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>gedrag</td>
<td>behavior</td>
<td>comportamiento</td>
<td>2.16</td>
<td>NA</td>
<td>2.09</td>
<td>6</td>
<td>8</td>
<td>14</td>
<td>gadras</td>
<td>possible cognate</td>
</tr>
<tr>
<td>geel</td>
<td>yellow</td>
<td>amarillo</td>
<td>1.08</td>
<td>1.72</td>
<td>1.94</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>keem</td>
<td>noncognate</td>
</tr>
<tr>
<td>geheim</td>
<td>secret</td>
<td>secreto</td>
<td>1.49</td>
<td>1.76</td>
<td>2.23</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>beheob</td>
<td>cognate</td>
</tr>
<tr>
<td>geld</td>
<td>money</td>
<td>dinero</td>
<td>2.44</td>
<td>2.61</td>
<td>2.60</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>gart</td>
<td>noncognate</td>
</tr>
<tr>
<td>geluid</td>
<td>sound</td>
<td>sonido</td>
<td>2.02</td>
<td>2.09</td>
<td>2.05</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>gemoil</td>
<td>cognate</td>
</tr>
<tr>
<td>geur</td>
<td>smell</td>
<td>olor</td>
<td>1.76</td>
<td>1.63</td>
<td>2.39</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>leop</td>
<td>possible cognate</td>
</tr>
<tr>
<td>geval</td>
<td>case</td>
<td>caso</td>
<td>2.61</td>
<td>2.58</td>
<td>2.98</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>terat</td>
<td>cognate</td>
</tr>
<tr>
<td>gevolg</td>
<td>result</td>
<td>consecuencia</td>
<td>2.18</td>
<td>2.19</td>
<td>2.30</td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>getang</td>
<td>possible cognate</td>
</tr>
<tr>
<td>gezicht</td>
<td>face</td>
<td>cara</td>
<td>2.65</td>
<td>2.60</td>
<td>2.80</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>dedight</td>
<td>noncognate</td>
</tr>
<tr>
<td>glimlach</td>
<td>smile</td>
<td>sonreir</td>
<td>1.88</td>
<td>1.87</td>
<td>1.79</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>blimack</td>
<td>noncognate</td>
</tr>
<tr>
<td>grap</td>
<td>joke</td>
<td>broma</td>
<td>1.18</td>
<td>1.52</td>
<td>1.74</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>kran</td>
<td>noncognate</td>
</tr>
<tr>
<td>heks</td>
<td>witch</td>
<td>bruja</td>
<td>1.04</td>
<td>1.20</td>
<td>1.45</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>wids</td>
<td>noncognate</td>
</tr>
<tr>
<td>hemd</td>
<td>shirt</td>
<td>camisa</td>
<td>1.32</td>
<td>1.65</td>
<td>2.08</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>tefs</td>
<td>noncognate</td>
</tr>
<tr>
<td>herfst</td>
<td>autumn</td>
<td>otoño</td>
<td>1.34</td>
<td>1.53</td>
<td>1.90</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>lartst</td>
<td>noncognate</td>
</tr>
<tr>
<td>huid</td>
<td>skin</td>
<td>piel</td>
<td>1.92</td>
<td>1.95</td>
<td>2.51</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>puit</td>
<td>noncognate</td>
</tr>
<tr>
<td>huwelijk</td>
<td>marriage</td>
<td>matrimonio</td>
<td>1.96</td>
<td>1.97</td>
<td>2.21</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>hikelijk</td>
<td>noncognate</td>
</tr>
<tr>
<td>kaak</td>
<td>jaw</td>
<td>mandíbula</td>
<td>0.95</td>
<td>1.08</td>
<td>1.20</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>zauk</td>
<td>possible cognate</td>
</tr>
<tr>
<td>kalk</td>
<td>lime</td>
<td>lima</td>
<td>1.04</td>
<td>0.90</td>
<td>0.00</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>vall</td>
<td>cognate</td>
</tr>
<tr>
<td>keus</td>
<td>choice</td>
<td>elección</td>
<td>1.48</td>
<td>2.01</td>
<td>1.84</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>waus</td>
<td>possible cognate</td>
</tr>
<tr>
<td>kleren</td>
<td>clothes</td>
<td>ropa</td>
<td>1.62</td>
<td>2.09</td>
<td>2.25</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>plimen</td>
<td>noncognate</td>
</tr>
<tr>
<td>kogel</td>
<td>ball</td>
<td>bala</td>
<td>1.20</td>
<td>1.97</td>
<td>1.34</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>kagep</td>
<td>possible cognate</td>
</tr>
</tbody>
</table>

* Dutch and English log frequencies were obtained via the WebCelex database, available online at [http://celex.mpi.nl](http://celex.mpi.nl) (Baayen, Piepenbrock, & Gulikers, 1995).
† Spanish log frequencies were obtained via the *Normas e índices de interés en Psicología Experimental* (NIPE, ‘Norms and Indices of interest in Experimental Psychology’) website (Díez, Fernández, & Alonso, 2006), which are taken from the Alameda & Cuetos (1995) corpus.
<table>
<thead>
<tr>
<th>Dutch word</th>
<th>English translation</th>
<th>Spanish translation</th>
<th>Dutch log frequency</th>
<th>English log frequency</th>
<th>Spanish log frequency</th>
<th>Dutch Length</th>
<th>English Length</th>
<th>Spanish Length</th>
<th>Dutch-like Nonword</th>
<th>English/Spanish Cognate Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>kooi</td>
<td>cage</td>
<td>jaula</td>
<td>1.26</td>
<td>1.11</td>
<td>1.49</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>soei</td>
<td>possible cognate</td>
</tr>
<tr>
<td>kring</td>
<td>circle</td>
<td>circulo</td>
<td>1.68</td>
<td>1.68</td>
<td>1.77</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>krajm</td>
<td>cognate</td>
</tr>
<tr>
<td>kwade</td>
<td>angry</td>
<td>enojado</td>
<td>1.23</td>
<td>1.81</td>
<td>0.48</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>gwice</td>
<td>noncognate</td>
</tr>
<tr>
<td>leeftijd</td>
<td>age</td>
<td>edad</td>
<td>2.00</td>
<td>2.39</td>
<td>2.60</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>peifigijd</td>
<td>noncognate</td>
</tr>
<tr>
<td>lepel</td>
<td>spoon</td>
<td>cucharra</td>
<td>1.04</td>
<td>1.04</td>
<td>0.85</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>tepol</td>
<td>noncognate</td>
</tr>
<tr>
<td>lezen</td>
<td>read</td>
<td>leer</td>
<td>1.88</td>
<td>1.51</td>
<td>2.29</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>feihan</td>
<td>noncognate</td>
</tr>
<tr>
<td>lichaam</td>
<td>body</td>
<td>cuerpo</td>
<td>2.42</td>
<td>2.47</td>
<td>3.00</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>bichaap</td>
<td>noncognate</td>
</tr>
<tr>
<td>mouw</td>
<td>sleeve</td>
<td>manga</td>
<td>1.15</td>
<td>0.95</td>
<td>1.53</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>houb</td>
<td>noncognate</td>
</tr>
<tr>
<td>neef</td>
<td>cousin</td>
<td>primo</td>
<td>1.28</td>
<td>1.34</td>
<td>1.71</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>ceep</td>
<td>noncognate</td>
</tr>
<tr>
<td>ogen</td>
<td>eyes</td>
<td>ojos</td>
<td>2.81</td>
<td>2.60</td>
<td>3.17</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>aben</td>
<td>noncognate</td>
</tr>
<tr>
<td>omgeving</td>
<td>area</td>
<td>ambiente</td>
<td>2.11</td>
<td>2.30</td>
<td>2.03</td>
<td>8</td>
<td>4</td>
<td>8</td>
<td>amwevang</td>
<td>possible cognate</td>
</tr>
<tr>
<td>oorzaak</td>
<td>cause</td>
<td>causa</td>
<td>1.90</td>
<td>1.87</td>
<td>2.42</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>oomdaik</td>
<td>cognate</td>
</tr>
<tr>
<td>ouder</td>
<td>parent</td>
<td>padre</td>
<td>1.78</td>
<td>1.85</td>
<td>3.01</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>aukem</td>
<td>possible cognate</td>
</tr>
<tr>
<td>paard</td>
<td>horse</td>
<td>caballo</td>
<td>2.00</td>
<td>1.93</td>
<td>2.27</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>kaarw</td>
<td>noncognate</td>
</tr>
<tr>
<td>poging</td>
<td>attempt</td>
<td>intento</td>
<td>1.81</td>
<td>1.88</td>
<td>1.99</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>pegang</td>
<td>possible cognate</td>
</tr>
<tr>
<td>rede</td>
<td>speech</td>
<td>discurso</td>
<td>1.58</td>
<td>1.88</td>
<td>2.26</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>vedu</td>
<td>possible cognate</td>
</tr>
<tr>
<td>reus</td>
<td>giant</td>
<td>gigante</td>
<td>1.08</td>
<td>1.54</td>
<td>1.53</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>riuk</td>
<td>possible cognate</td>
</tr>
<tr>
<td>ruim</td>
<td>wide</td>
<td>espaciado</td>
<td>1.85</td>
<td>1.91</td>
<td>0.30</td>
<td>4</td>
<td>4</td>
<td>9</td>
<td>Jame</td>
<td>possible cognate</td>
</tr>
<tr>
<td>steun</td>
<td>support</td>
<td>apoyo</td>
<td>1.76</td>
<td>2.12</td>
<td>1.83</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>steid</td>
<td>noncognate</td>
</tr>
<tr>
<td>stevig</td>
<td>firm</td>
<td>firme</td>
<td>1.64</td>
<td>1.58</td>
<td>2.02</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>stilag</td>
<td>cognate</td>
</tr>
<tr>
<td>strak</td>
<td>tight</td>
<td>apretado</td>
<td>1.62</td>
<td>1.53</td>
<td>0.95</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>sthod</td>
<td>noncognate</td>
</tr>
<tr>
<td>tijd</td>
<td>time</td>
<td>tiempo</td>
<td>2.98</td>
<td>3.25</td>
<td>3.35</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>dijl</td>
<td>cognate</td>
</tr>
<tr>
<td>touw</td>
<td>rope</td>
<td>cuerda</td>
<td>1.40</td>
<td>1.49</td>
<td>1.69</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>roag</td>
<td>possible cognate</td>
</tr>
<tr>
<td>trage</td>
<td>slow</td>
<td>lento</td>
<td>1.08</td>
<td>1.78</td>
<td>1.79</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>fraga</td>
<td>noncognate</td>
</tr>
<tr>
<td>trui</td>
<td>sweater</td>
<td>sueter</td>
<td>1.15</td>
<td>1.04</td>
<td>0.78</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>kruo</td>
<td>cognate</td>
</tr>
<tr>
<td>tuin</td>
<td>garden</td>
<td>jardín</td>
<td>1.99</td>
<td>2.04</td>
<td>2.27</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>meib</td>
<td>noncognate</td>
</tr>
<tr>
<td>Dutch word</td>
<td>English translation</td>
<td>Spanish translation</td>
<td>Dutch log frequency</td>
<td>English log frequency</td>
<td>Spanish log frequency</td>
<td>Dutch Length</td>
<td>English Length</td>
<td>Spanish Length</td>
<td>Dutch-like Nonword</td>
<td>English/Spanish Cognate Status</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>verf</td>
<td>paint</td>
<td>pintar</td>
<td>1.42</td>
<td>1.38</td>
<td>1.51</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>mert</td>
<td>cognate</td>
</tr>
<tr>
<td>verhalen</td>
<td>stories</td>
<td>cuentos</td>
<td>1.84</td>
<td>1.79</td>
<td>1.85</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>merhogen</td>
<td>noncognate</td>
</tr>
<tr>
<td>vijand</td>
<td>enemy</td>
<td>enemigo</td>
<td>1.60</td>
<td>1.70</td>
<td>2.01</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>rijand</td>
<td>cognate</td>
</tr>
<tr>
<td>vlak</td>
<td>plane</td>
<td>plano</td>
<td>1.61</td>
<td>1.65</td>
<td>2.07</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>vlop</td>
<td>cognate</td>
</tr>
<tr>
<td>vlieg</td>
<td>fly</td>
<td>volar</td>
<td>0.95</td>
<td>1.23</td>
<td>1.72</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>zliug</td>
<td>noncognate</td>
</tr>
<tr>
<td>vlug</td>
<td>fast</td>
<td>rápido</td>
<td>1.91</td>
<td>1.84</td>
<td>1.98</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>milg</td>
<td>possible cognate</td>
</tr>
<tr>
<td>voogd</td>
<td>tutor</td>
<td>tutor</td>
<td>1.26</td>
<td>1.23</td>
<td>0.85</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>doehd</td>
<td>cognate</td>
</tr>
<tr>
<td>vraag</td>
<td>question</td>
<td>pregunta</td>
<td>2.53</td>
<td>2.45</td>
<td>2.40</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>traef</td>
<td>noncognate</td>
</tr>
<tr>
<td>vrede</td>
<td>peace</td>
<td>paz</td>
<td>1.71</td>
<td>1.95</td>
<td>2.47</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>vlude</td>
<td>noncognate</td>
</tr>
<tr>
<td>vrijheid</td>
<td>freedom</td>
<td>libertad</td>
<td>2.08</td>
<td>2.00</td>
<td>2.57</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>prijbied</td>
<td>possible cognate</td>
</tr>
<tr>
<td>vrouw</td>
<td>woman</td>
<td>mujer</td>
<td>2.78</td>
<td>2.53</td>
<td>3.13</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>vreik</td>
<td>noncognate</td>
</tr>
<tr>
<td>wens</td>
<td>wish</td>
<td>deseo</td>
<td>1.58</td>
<td>1.45</td>
<td>2.38</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>vels</td>
<td>possible cognate</td>
</tr>
<tr>
<td>wijk</td>
<td>district</td>
<td>barrio</td>
<td>1.15</td>
<td>1.58</td>
<td>2.21</td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>rijk</td>
<td>possible cognate</td>
</tr>
<tr>
<td>winkel</td>
<td>store</td>
<td>tienda</td>
<td>1.57</td>
<td>1.69</td>
<td>1.96</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>liggel</td>
<td>noncognate</td>
</tr>
<tr>
<td>zake</td>
<td>respect</td>
<td>respeto</td>
<td>1.26</td>
<td>1.77</td>
<td>2.15</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>tabe</td>
<td>cognate</td>
</tr>
<tr>
<td>zeep</td>
<td>soap</td>
<td>jabón</td>
<td>1.20</td>
<td>1.30</td>
<td>1.26</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>peeg</td>
<td>noncognate</td>
</tr>
<tr>
<td>zegen</td>
<td>blessing</td>
<td>bendición</td>
<td>1.04</td>
<td>1.00</td>
<td>1.20</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>bigeg</td>
<td>possible cognate</td>
</tr>
<tr>
<td>zoek</td>
<td>search</td>
<td>buscar</td>
<td>1.52</td>
<td>1.67</td>
<td>2.28</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>naek</td>
<td>noncognate</td>
</tr>
<tr>
<td>zorgen</td>
<td>care</td>
<td>preocupación‡</td>
<td>1.71</td>
<td>1.97</td>
<td>1.73</td>
<td>6</td>
<td>4</td>
<td>14</td>
<td>diken</td>
<td>possible cognate</td>
</tr>
<tr>
<td>zout</td>
<td>salt</td>
<td>sal</td>
<td>1.54</td>
<td>1.54</td>
<td>1.78</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>poun</td>
<td>cognate</td>
</tr>
<tr>
<td>zure</td>
<td>acid</td>
<td>ácido</td>
<td>1.04</td>
<td>1.18</td>
<td>1.74</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>zeri</td>
<td>cognate</td>
</tr>
<tr>
<td>zwaar</td>
<td>heavy</td>
<td>pesado</td>
<td>1.96</td>
<td>2.12</td>
<td>1.57</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>hwoer</td>
<td>noncognate</td>
</tr>
<tr>
<td>zweeg</td>
<td>silent</td>
<td>silencioso</td>
<td>1.62</td>
<td>1.79</td>
<td>1.66</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>fraeg</td>
<td>cognate</td>
</tr>
</tbody>
</table>
CARI ANNE BOGULSKI

Education
Pennsylvania State University, University Park, PA  Ph.D.  December 2013
Pennsylvania State University, University Park, PA  M.S.  December 2007
University of Arkansas, Fayetteville, AR  B.A.  May 2007

Awards
National Science Foundation Doctoral Dissertation Research Grant  2011-2013
Penn State PIRE Graduate Fellowship, Penn State University  2011
RGSO Dissertation Grant, Penn State University  2011
Adele Miccio Memorial Travel Award, CLS, Penn State University  2012
National Science Foundation Graduate Research Fellowship  2008-2011
University Graduate Fellowship  2008
Yoder Fellowship  2007

Publications and Papers in Preparation
Bogulski, C. A., & Kroll, J. F. (in preparation). When bilingualism incurs a cost to language production: Weaker links or cross-language competition?


