THE EFFECTS OF PROFICIENCY LEVEL ON LEXICAL ACCESS:
AN ELECTROPHYSIOLOGICAL INVESTIGATION

A Thesis in
Communication Sciences and Disorders

by

Elizabeth A. Willis

© 2011 Elizabeth A. Willis

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2011
The thesis of Elizabeth A. Willis was reviewed and approved* by the following:

Maya Misra  
Assistant Professor of Communication Sciences and Disorders  
Thesis Adviser

Carol A. Miller  
Associate Professor of Communication Sciences and Disorders

Kathryn D.R. Drager  
Associate Professor of Communication Sciences and Disorders

*Signatures are on file in the Graduate School
ABSTRACT

This study assessed how lexical access changes as a result of proficiency level by examining the performance of two groups of English-Spanish bilinguals, a “less proficient” and a “more proficient” group, in a task that asked them to decide whether English words were the correct translation of Spanish words. Critical conditions were trials on which incorrect translations were related in lexical form (“translation related distractors”) or meaning (“semantic distractors”) to the correct translation. Behavioral and ERP measures revealed interference for both distractor types in both groups. Both types of distractor conditions produced slower and less accurate responses than their matched controls. However, ERPs revealed a different time-course of effects and locus of interference based on proficiency level, specifically for the translation related distractors. In both groups the semantically related distractors attenuated the N400 component of the ERPs, which is known to be sensitive to lexical and semantic access. In the translation related trials, the more proficient group showed evidence for an inhibitory process, while the less proficient group showed evidence for a later process associated with conscious checking of the item before a decision was made. Overall, results provided evidence that lexical and conceptual information are active regardless of a bilingual’s proficiency level, but the differences in the ERP data between groups suggests that the way in which lexical and conceptual information is utilized varies as a function of L2 skill. Specifically, it is hypothesized that the less proficient bilinguals may have considered English words related in lexical form as potential translations, while more proficient bilinguals may have found it difficult to process these distractors because they had already inhibited them. Since effects for words related in meaning were more similar between groups, the
processes underlying conceptual access may be similar in second language learners and skilled bilinguals.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................ vi

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

The Revised Hierarchical Model for Bilingual Lexical Access ................................. 1
Evidence for Access to Concepts in Early L2 Learners ............................................. 4
Evidence for Continued Access to Translation Equivalents in Skilled Bilinguals .................................................................................................................. 7
Event-Related Potential Technique ........................................................................... 8
ERP Studies of Lexical Access in Bilinguals ............................................................... 12

The Current Study .................................................................................................... 15

Methods ................................................................................................................... 16

Participants ............................................................................................................. 16
Experimental Overview ......................................................................................... 20
Stimuli .................................................................................................................... 22
EEG Recording ....................................................................................................... 27
Data Analysis .......................................................................................................... 28
  Behavioral data analysis ....................................................................................... 28
  ERP data analysis ................................................................................................. 29

Results .................................................................................................................... 31

Behavioral data ...................................................................................................... 31
Visual Inspection of ERP Data .................................................................................. 33
  All Participants .................................................................................................... 33
  Less Proficient Group ......................................................................................... 35
  More Proficient Group ....................................................................................... 36
Statistical Analysis of ERP Data ............................................................................... 40
  Translation Interference Effect ........................................................................... 40
    P200 .................................................................................................................. 40
    N400 ............................................................................................................... 40
    LPC .................................................................................................................. 42
  Semantic Interference Effect .............................................................................. 43
    P200 .................................................................................................................. 43
    N400 ............................................................................................................... 44
    LPC .................................................................................................................. 44

Discussion ............................................................................................................... 45

References .............................................................................................................. 52
LIST OF TABLES

Table 1. Participant Characteristics.........................................................19
Table 2. Sample stimuli for yes trials and critical conditions used in Task 1.........25
Table 3. Average word length and KFFRQ for each stimulus list (order) across critical conditions in Task 1..............................................................26
LIST OF FIGURES

Figure 1. Translation recognition task: Time course and organization of stimuli from the onset of one critical item to the next .................................................................21

Figure 2. 32-channel electrode montage ...............................................................28

Figure 3. Behavioral interference for mean reaction times (in milliseconds) for translation related distractors versus translation controls and semantically related distractors versus semantic controls for the less proficient and more proficient groups ........................................32

Figure 4. Behavioral interference for percent accuracy for translation related distractors versus translation controls and semantically related distractors versus semantic controls for the less proficient and more proficient groups ...........................................33

Figure 5. Grand average ERPs elicited by translation related distractors and translation controls for all 24 participants. .................................................................34

Figure 6. Grand average ERPs elicited by semantically related distractors and semantic controls for all 24 participants. .................................................................34

Figure 7. Grand average ERPs elicited by translation related distractors and translation controls for the less proficient group (n=12). .................................................36

Figure 8. Grand average ERPs elicited by semantically related distractors and semantic controls for the less proficient group (n=12). .................................................36

Figure 9. Grand average ERPs elicited by translation related distractors and translation controls for the more proficient group (n=12) ...............................................38

Figure 10. Grand average ERPs elicited by semantically related distractors and semantic controls for the more proficient group (n=12) ...............................................38

Figure 11. Grand average ERPs for translation related distractors and translation controls
for (A) the less proficient group and (B) the more proficient group at three midline sites, Fz, Cz, and Pz.

Figure 12. Grand average ERPs for semantically related distractors and semantic controls for (A) the less proficient group and (B) the more proficient group at three midline sites, Fz, Cz, and Pz.
When learning a first language, it seems clear that individuals learn to associate words with concepts. However, a fundamental question in research on bilingualism is whether individuals access words in their second language (L2) the same way as native speakers or if there are differences in the process, with learners associating words in the second language with words in their native language rather than to concepts directly. Researchers have debated how second language learners access the meaning of words in their L2 and whether proficiency impacts their approach. One of the most widely accepted models of bilingual lexical representation is the Revised Hierarchical Model (RHM), which will be discussed in detail below. While there is evidence to suggest that certain aspects of this model are true, several studies have found results that are in conflict with its claims. The aim of the current study is to test the claims made by the RHM, as well as the conclusions drawn by subsequent studies in order to establish a more consistent developmental trajectory for lexical access as it relates to proficiency level.

*The Revised Hierarchical Model for Bilingual Lexical Access*

Kroll and Stewart’s (1994) RHM suggests that, to activate the meaning of L2 words, less proficient learners rely on lexical links with translation equivalents in their L1, but more proficient learners access concepts more directly from L2 words. In other words, it is hypothesized that the reliance on the L1 to access words in the L2 is a function of the degree of skill in the L2. Not surprisingly, the link between words and concepts (conceptual link) is thought to be stronger for L1 than for L2; a strong established link between words in L1 and conceptual memory would be expected to be formed during early childhood when the first language is initially learned. During early second language learning for sequential bilinguals, it is hypothesized that L2 words come
into the system via lexical links with words in the established L1. Thus, initially a second language word would be associated with its translation (perhaps because of the common instructional approach in L2 classrooms of teaching words in relation to their translations), and it is the L1 translation that is directly linked to meaning. As proficiency increases, and the developing bilingual has more experience with using the words in the L2 referentially, the L2 can establish its own direct conceptual links. It is important to note that the lexical link remains, and that the L1 link to conceptual memory is still thought to be stronger than the new L2 conceptual link. However, over time, the L2 conceptual link can strengthen. Therefore, a central implication of the RHM is that with increased proficiency there is a developmental shift from the L2’s reliance on L1 to access the meaning of words to a more direct access to concepts resulting from the shared conceptual system.

A number of studies have tested the claims of the RHM, and several of them have produced results that are consistent with its predictions. For example, Talamas, Kroll, & Dufour (1999) directly examined the developmental progression predicted by the RHM (i.e., that with increased proficiency, bilinguals move from a reliance on lexical form to a reliance on meaning, Kroll & Stewart, 1994) using a translation recognition task. In a translation recognition task, participants must decide whether the second word in a pair is the correct translation of the first word (De Groot, 1992). Many studies have utilized this paradigm to evaluate processes underlying bilingual lexical access, including the current study. Talamas et al. found that relatively skilled English-Spanish bilinguals were more sensitive to interference from L1 distractors related in meaning than in form to the L2 words, whereas the opposite was true for less skilled learners. For example, more skilled
Spanish learners were slower and less accurate to reject incorrect pairs such as “pecho” – “rib” (semantically related to the correct translation, “chest”) than pairs such as “pecho” – “chess” (related in form to the correct translation), while less skilled participants showed greater interference for the form-related distractors. The fact that the more proficient participants were slower in responding to semantically related items is indicative of a processing cost due to concept mediation, while the interference from distractors related in form to the correct translation in the less proficient group revealed lexical competition. Another important finding in the Talamas et al. study comes from the analysis of task accuracy data. The results here offer mixed support for the influence of conceptual processing in the early stages of second language learning. While the less proficient group did not show significant effects in reaction time for semantically related words, task accuracy was compromised, revealing that semantic information did in fact play a role in processing even in the less proficient group. These results support the idea that concept mediation progresses gradually as a function of proficiency level (Kroll & Stewart, 1994). In other words, as proficiency increases, so does the influence of meaning on lexical access, although some sensitivity to conceptual information may exist even from the early stages of L2 acquisition. In addition, results suggest that reliance on direct lexical links between words in the L1 and L2 diminishes with proficiency, as less interference from items related in form to the correct translations was found for the more proficient group.

Sunderman and Kroll (2006) also utilized a translation recognition task to investigate lexical processing in English-Spanish bilinguals with two groups of English-Spanish participants, one less proficient and one more proficient in Spanish, but their
study introduced some refinements over the Talamas et al. (1999) paradigm. Sunderman and Kroll included distractors for the following conditions: words that were related in form to the Spanish item (lexical neighbors; e.g., “cara” – “card”, where the correct translation for “cara” is “face”), words that were related in form to the correct translation equivalent (translation neighbors; e.g., “cara” – “fact”), or words that were related in meaning to the correct translation (semantically related; e.g., “cara” – “head”). Sunderman and Kroll reported translation related form interference for translation neighbors only for less proficient L2 learners while semantic interference and form interference for lexical neighbors were seen in both less and more proficient groups. As with the earlier study by Talamas et al., the more proficient group did not appear to rely on L1 mediation to understand the meaning of L2 words, thus supporting one of the claims of the RHM. However, this experiment also provided further evidence that information about meaning is accessible even during early L2 learning, when a reliance on lexical information may be assumed under the RHM (Sunderman & Kroll).

Evidence for Access to Concepts in Early L2 Learners

The studies by Talamas et al. (1999) and Sunderman and Kroll (2006) are consistent with the RHM’s prediction that as learners become more proficient in their L2, they are less likely to utilize L1 mediation to understand words. Several studies even suggest that L1 mediation is not required, even for individuals at the earliest stages of L2 learning (e.g., Altarriba & Mathis, 1997; Dufour & Kroll, 1995; Sunderman & Kroll). In other words, learners have immediate access to the concepts associated with L2 words. A study by Dufour and Kroll investigated the theory of concept mediation and bilingual representation in two groups of English-French bilinguals. Specifically, the researchers
utilized a categorization paradigm to examine the interactions between conceptual and lexical links, and the effect that proficiency has on these connections. The task asked participants to view category names and then decide whether a target item was a member of that category. Category names and target items appeared in either English or French across experimental conditions. Based on past research findings, including implications of the RHM, the experimenters hypothesized that less proficient participants would rely on lexical mediation or word association techniques to translate L2 words, while more proficient participants would be able to directly access meanings. Following this assumption, the more proficient bilinguals were expected to be faster and more accurate at the task of semantic categorization. Furthermore, Dufour and Kroll argued that if participants are able to directly access the concept from the L2 word, i.e., use concept mediation, then the language of the category name should not affect performance. Conversely, if lexical mediation is required, participants should experience more difficulty in the form of longer response latencies for categories and items presented in the L2 than in the L1.

Dufour and Kroll’s (1995) results showed that the response latency for more proficient bilinguals was not affected by the language of the category name, suggesting that they were able to conceptually mediate in both languages and that both languages were accessing a shared conceptual representation. However, the less proficient participants were affected significantly by the language of the target items and the category names. The less proficient group was faster at categorizing when the language of the target items matched that of the category name (regardless of the language tested) compared to when the language of the target item and category name was different,
suggesting that access to conceptual information from their L2 was limited in the less proficient group. This result also suggests that the less proficient group did not appear to rely on the translation of L2 words into L1 in order to access concepts. If the less proficient participants had been translating French words into English in order to perform the task, then the French – French condition would have produced the longest response latencies. The researchers interpreted the results to mean that while both groups of bilinguals utilized conceptual links, the less proficient group had weaker conceptual links that were only present in categorization within the same language. The finding that the connection between L2 words and concepts was present, but weaker for less proficient bilinguals, supports the idea that bilingual memory and lexical representation change as a function of L2 proficiency, which is a concept that is in concordance with the RHM.

Another study by Altarriba and Mathis (1997) examined the differences in lexical and conceptual representation in second language acquisition through the use of translation recognition tasks. The aim of the study was to directly investigate Kroll and Stewart’s (1994) model of bilingual memory representation by testing the model with novice second language learners. The experimenters highly trained these learners in the Spanish to English translations that were later tested in the tasks. This study is different than the other studies mentioned in that it included a group of monolingual participants learning a language for the first time. Specifically, monolingual English learners of Spanish were included, and they were compared to a second group of “expert” bilinguals. In their first experiment, Altarriba and Mathis found interference for words that were orthographically related to the correct translations in both groups, indicating the presence of shared lexical level links in all participants. An interference effect for semantically
related distractors was also found in both groups in a second experiment. In their third experiment, a Stroop task was used to identify the presence of conceptual interference. In a Stroop task participants view color words (e.g., red or green) written in different ink colors and are asked to name the colors of the ink. The bilingual version has the color names written in either the L1 or L2 across trials and asks participants to name the colors of the ink in either the L1 or L2. The researchers hypothesized that only fluent bilinguals would demonstrate a between-language Stroop effect since the effect is a result of interference at the conceptual level. However, between-language Stroop effects were found in both the expert and novice groups. The authors claimed that the fact that a Stroop effect was apparent for the learners suggests that they used concept mediation in the translation of L2 words. Altarriba and Mathis interpreted their results to imply that L1 mediation was not required, even for individuals in the earliest stages of L2 acquisition. As a result, Altarriba and Mathis argued that the RHM is wrong in certain aspects. They claimed that it is the specific words known in the L2, and not the proficiency level, that should drive the model. However, the conclusions drawn from this experiment are questionable due to the nature of the group of learners. One can argue that this group should not be compared to groups of “less proficient bilinguals” in other studies, as the effect of their intense vocabulary training is unclear.

**Evidence for Continued Access to Translation Equivalents in Skilled Bilinguals**

An important finding in the study by Altarriba and Mathis (1997) was the presence of orthographic effects not only in the novice group, but in the “expert” group, as well. This finding is consistent with a more recent line of research suggesting that even after a high degree of skill has been acquired, proficient bilinguals may continue to access the
translation equivalent (e.g., Guasch, Sánchez-Casas, Ferré, and García-Albea, 2008; Guo, Misra, Tam, & Kroll, submitted; Thierry and Wu, 2007). For example, a study by Guasch et al. asked beginning and intermediate Spanish-Catalan learners and highly proficient bilinguals to perform a translation recognition task. There were three types of word types in the experiment: very closely semantically related word pairs, closely related word pairs, and form-related pairs. The results showed an effect for semantically related word pairs only in the highly proficient group, indicating that the direct access to semantic information was dependent on the participants’ level of proficiency. More surprisingly, results also showed an effect for the form manipulation in all three participant groups, indicating that even the highly proficient bilinguals were utilizing information at the lexical level. These results thus suggest that the translation equivalent may still be activated in bilinguals with a high degree of skill.

Two other recent studies utilizing event-related potentials (ERPs) have also shown evidence for activation of translation equivalents in skilled bilinguals (Guo et al., submitted; Thierry & Wu, 2007). These recent studies will be discussed after a brief review of the ERP method.

*Event-Related Potential Technique*

ERPs are an online method that can be used to evaluate the neurocognitive correlates of language processing with excellent temporal precision, and recent experiments with L2 learners have shown that ERPs may be sensitive to processing that is obscured in the behavioral record (e.g., McLaughlin, Osterhout, & Kim, 2004; Tokowicz & MacWhinney, 2005). Therefore, it is an excellent candidate method for disentangling the time-course of lexical access in bilinguals.
The ERP technique is a derivative of electroencephalography (EEG), which involves the measurement of the brain’s spontaneous electrical activity as recorded by multiple electrodes placed on the scalp (see Molfese, Molfese, & Kelly, 2001, for a review). The electrical activity measured in EEG is the summation of the synchronous activity (neural firing) of spatially related cortical neurons. ERPs focus on a portion of the ongoing electrical activity presented in the EEG by extracting neural responses to motor, sensory, or cognitive stimuli. That is, the neural responses elicited are time-locked to the onset of the stimuli, so that it is possible to analyze the effects of the stimuli and change in the EEG pattern. This allows researchers to assess the temporal relationship between the neuroelectrical responses and the given event (Molfese et al.). ERPs are extracted from the ongoing EEG by means of filtering and signal averaging and results are generally evaluated in the time domain, as waveforms that plot the change in voltage over time (Picton et al., 2000). They are sampled in milliseconds (ms), which accounts for their strong temporal resolution and accuracy in the measurement of information processing in the brain. The spatial resolution is more limited, but multichannel recordings can allow researchers to estimate the intracerebral locations of the processes that result from the presentation of various stimuli (Picton et al.), and even without this information the topography of the voltage changes can be informative with respect to the similarity or differences between waveforms.

Visual analysis of the ERP waveform involves the identification of positive and negative peaks, called components. A labeling system is used to indicate both the sequence in which a peak occurs and its polarity (e.g., N2 would refer to the second negative peak in the waveform). Peaks may also be named based on their latency, which
is defined as the amount of time elapsed from the onset of the stimulus (e.g., P600 is a positive peak occurring 600 ms post-stimulus onset) (Molfese et al., 2001). The analysis of various components in the ERP waveform provides insight into how individual participants process information. Different ERP components reflect aspects of online information processing in the brain, and three components are of specific interest to the current work and prior studies evaluating lexical access: the P200, the N400, and the late positive component (LPC).

The P200 is a component in the ERP waveform that has proven to be sensitive to lexical processing in previous studies; however, results have been mixed regarding whether the effect is demonstrated by an increase or reduction in amplitude. For example, Barnea and Breznitz (1998) found an increase in P200 amplitude for word pairs that rhymed in their study investigating orthographic and phonological processing. They suggested that the P200 component serves as an early index of word recognition. Misra and Holcomb (2003) also found an enhanced P200 for masked repetition priming in a semantic monitoring task. Conversely, a reduction in P200 amplitude was found for orthographically related Chinese word pairs in a study by Liu, Perfetti, and Hart (2003), and in a version of Thierry and Wu’s (2007) semantic judgment task in which native Chinese speakers read Chinese word pairs that included one repeated character. While results have been inconsistent as to whether an increase or reduction in P200 amplitude is observed, the component is recognized as one that is sensitive to form similarity between two words and can provide important information on the time-course of lexical processing.
The second ERP component of interest to the current study is the N400. The N400 was first observed by Kutas and Hillyard (1980), who presented readers with sentences in which the final word of the sentence was either congruous or incongruous. They found that incongruous words elicited a larger negative ERP component that peaked approximately 400 ms following the onset of the word, suggesting that this component indexes semantic integration processes. N400s can also be generated by any word presented in isolation (i.e., in a simple word reading experiment) and can be impacted by a single context word. For example, the N400 for the word “doctor,” presented soon after the word “nurse” would be smaller than the N400 to the word “doctor” presented after the word “table” (e.g., Rugg, 1985). N400s have been shown to be sensitive to lexical and semantic manipulations. As previously mentioned, the N400 typically shows attenuated amplitudes for items that are easier to retrieve or integrate, either due to familiarity (e.g., with repetition), or context (e.g., due to sentence context or even the prior activation of a single semantically related word) (Misra & Holcomb, 2003).

According to Rugg (1990), a late positive component (LPC) is one that occurs slightly after the N400 and is usually larger for repeated items than it is for items that are presented for the first time in an experiment. It is hypothesized that the LPC reflects aspects of episodic memory retrieval (Van Petten et al., 1991), and the LPC is associated with experimental tasks that involve conscious recollection of the relationship between items (e.g., Misra & Holcomb, 2003). However, late positive components have also been observed in tasks which do not involve direct repetition. The LPC often overlaps with another component referred to as the P600, and some argue that the two are from the same family of components (Guo et al., submitted). Kolk and Chwilla (2007) have argued
that effects on the P600 are elicited as a result of reprocessing information and attempts at resolving response uncertainty. This hypothesis suggests that the LPC/P600 component reflects reanalysis or reprocessing of information.

**ERP Studies of Lexical Access in Bilinguals**

In their ERP study, Thierry and Wu (2007) asked Chinese-English bilinguals to perform a task in which they decided whether or not two English words were semantically related. The participants were unaware that when translated into Chinese, half of the trials featured words that included shared characters. In other words, the translations of the two words were related in lexical form to one another. This study was unique in that it assessed the influence of the L1 (Chinese) in a task that required participants to perform strictly in their L2 (English). Thierry and Wu utilized both ERPs and behavioral measures to determine whether the Chinese translations of the English words were activated during this task. While the behavioral data did not show sensitivity to the hidden Chinese characters, there was an effect evident in the ERP data, indicating that the Chinese translations were activated. Specifically, the ERPs showed an attenuated N400 for unrelated English word pairs with a shared character in their Chinese translations in comparison to unrelated words pairs without a shared character. These results provide further evidence that even relatively proficient bilinguals access L1 translations when reading words in their L2.

An experiment by Guo et al. (submitted) also tested whether relatively proficient Chinese-English bilinguals access translation equivalents while reading in L2, but it assessed behavioral and ERP effects on a translation recognition task. Critical trials were translations related in lexical form (“translation form distractors”) or meaning (“semantic
distractors”) to the correct translations. Their initial experiment revealed similar behavioral interference for both conditions, while ERPs showed a different time-course of effects based on the distractor type. Specifically, translation form distractors elicited a larger P200 and a larger late positive component (LPC) than unrelated controls, while the effects found for the semantic distractors were primarily located on the N400 and LPC, and the LPC effect for semantic distractors differed from that observed for translation form distractors. The P200 effect for the translation form distractors was hypothesized to reflect automatic sensitivity to the form similarity between the translation of the L2 word and the L1 word, while the N400 effect for the semantic distractors was interpreted to show that processing of the L2 word in this condition provided sufficient context to activate the semantically related L1 word. The LPC effect in both conditions was hypothesized to reflect conscious reprocessing or “checking” to verify that the L1 word was not, in fact, a translation of the L2 word. These results are consistent with those reported by Thierry and Wu (2007), that even for relatively proficient bilinguals, the L1 translation is activated while reading L2 words. However, Guo et al. argued that the L1 translation may have been activated after the concept in these relatively proficient bilinguals, rather than as a means to lexical access. Specifically, they showed in a second experiment, using behavioral measures only, that the stimulus onset asynchrony (SOA) between the L2 and L1 words impacted the pattern of results by testing a SOA of 300 milliseconds (ms) in addition to the 750 ms SOA that was used in their ERP study. The shorter SOA was more similar to that used in previous studies using the translation recognition task to support the RHM (e.g., Talamas et al., 1999; Sunderman & Kroll, 2006). As was expected, the 750 ms SOA in the second experiment produced the same
effects as in the first experiment. However, results for the 300 ms SOA were different in that only semantic interference was found in the RTs. Accuracy data at the short SOA provided evidence for both semantic and translation interference, although the effects for semantic interference were stronger. These researchers concluded that relatively proficient bilinguals can directly access the meaning of L2 words, and that it is not necessary to access the translation equivalent, although it may be activated as a consequence of an interactive lexical system. In other words, it may have been the increased length of the SOA in some past studies (750 ms in Guasch et al., 2008; 1100 ms in Thierry & Wu, 2007) that encouraged access to the translation equivalent.

It is clear that past research has produced conflicting results on the degree to which bilinguals rely on lexical versus conceptual links while accessing words in their second language. Some studies have suggested that as a bilingual’s level of proficiency increases, the bilingual progresses from a reliance on form to a reliance on meaning (e.g., Kroll & Stewart, 1994; Talamas et al., 1999). In other words, the more proficient bilingual will have the ability to directly access concepts in their L2, while the less proficient bilingual has minimal access to concepts initially and therefore relies on lexical links to access L2 words. Other studies have suggested that learners have immediate access to the concepts associated with L2 words, and that L1 mediation is not required at all (e.g., Altarriba & Mathis, 1997; Dufour & Kroll, 1995; Sunderman & Kroll, 2006). Finally, another line of research suggests that even highly proficient bilinguals may continue to access the translation equivalent while reading words in their second language (e.g., Guasch et al., 2008; Thierry and Wu, 2007), but that whether or not they access the translation equivalent may depend on the length of time that they have to
perform a task (Guo et al., submitted).

The Current Study

The current study takes another look at how lexical access changes as a result of proficiency level by examining the performance of two groups of English-Spanish bilinguals in a translation recognition task. Like the studies by Thierry and Wu (2007) and Guo et al. (submitted), the current study exploited the sensitivity of comparing event-related potentials (ERPs) and behavioral measures to evaluate the time-course of activation of lexical and conceptual information. However, while Thierry and Wu and Guo et al. tested relatively proficient Chinese-English bilinguals, the current study evaluated the developmental claims of the RHM by comparing more and less proficient English-Spanish bilinguals. This language pairing in this population is also more similar to that tested in other studies using the translation recognition task to evaluate the claims of the RHM (e.g., Guasch et al., 2008; Sunderman & Kroll, 2006; Talamas et al., 1999). Although similar patterns of results are often observed regardless of the language pairing (e.g., Hoshino & Kroll, 2008), the possibility still remains that results in Thierry and Wu and Guo et al. may have differed from prior reports due to the Chinese-English bilinguals’ script differences.

The bilinguals in the current study were first presented with a word in their L2 (Spanish) and then a word in their L1 (English). As with other translation recognition studies, the critical trials were ones in which the word pairs were not correct translations. Similar to the Guo et al. (submitted) study, “translation related distractors” included words that were related in lexical form to the correct translation (e.g., a translation distractor for the word “playa” which means “beach”, was “batch”), while “semantically
related distractors” included words that were related in meaning to the correct translations (e.g., the semantically related distractor for the word “playa” was “sand”). In addition to the two types of distractors, unrelated matched controls were included for each distractor condition. Also, unlike the Guo et al. study, the same L2 (Spanish) words were used in both types of critical distractor conditions and in the control conditions, better allowing for comparison of translation form and semantic effects across items. The goals of the current study were as follows:

1. To investigate the neural correlates of lexical access in bilingual English-Spanish speakers using a translation recognition task which requires participants to switch between their primary and secondary languages.

2. To determine whether the observed patterns on the translation recognition task vary with the participants' proficiency in their second language.

3. To assess whether differences between more and less proficient bilinguals appear to reflect changes in the reliance on lexical versus conceptual links, as a result of increasing proficiency.

Methods

Participants

Participants for the current study were English-Spanish bilinguals all recruited from the Pennsylvania State University community. Results are presented for 24 participants (21 females, 3 males), ages 18-27 (average age 22 ± 2.9 yrs). A total of 47 participants initially completed the experiment, but data from 23 participants were excluded from the current study. Data from five participants were excluded due to an insufficient number of trials, primarily due to unrecoverable eye blink artifacts. One participant’s data was not useable as a result of difficulties in lowering the impedances
during the electrode application process. Six participants failed to meet the accuracy criteria for the task (70% or greater for the “yes” translation trials), and the remaining 11 participants did not meet the language inclusion criteria for the current study (i.e., they were native Spanish speakers or simultaneous language learners). Data from these 11 participants will be summarized elsewhere. Participant characteristics for the included 24 subjects are described below and summarized in Table 1.

All participants were right-handed with no history of neurological or language disorders and had normal or corrected-to-normal vision and hearing. Participants in the final set were native English speakers who were also proficient in Spanish. They considered themselves sequential Spanish learners and all reported that English was their dominant language, as determined by self-ratings and responses on a language history questionnaire (LHQ). Specifically, the LHQ asked participants to rate their language skills in five domains (i.e., reading, spelling, writing, speaking, and comprehension) on a 10-point scale (1 = not literate/fluent, 10 = very literate/fluent). As a group, the participants’ ratings of their English (L1) skills were higher than their Spanish (L2) skills (L1: mean = 9.83, SD = 0.27; L2 mean = 7.26, SD = 1.32).

Since the goal of the current study was to assess the role of proficiency on lexical access, participants were split into two groups based on their proficiency level. The two groups were a “more proficient” group and a “less proficient” group. The primary criterion for group placement was study abroad experience. All participants with study abroad experience in a Spanish-speaking country were placed into the “more proficient” group. In a review of L2 learning in study abroad contexts, Freed (1998) suggested that there were significant linguistic benefits for these learners. For example, they tend to
speak with greater ease and confidence and tend to have a greater abundance of speech and enhanced fluency overall. Additionally, Milton and Meara (1995) reported that a study abroad context is an important agent in the acquisition of vocabulary. The range of time spent abroad for the participants in the current study was 1 month to 27 months, with an average length of 9.38 months.

In order to assess whether group placement based on study abroad experience was appropriate, self-ratings from the language history questionnaires and results from a lexical decision task (LDT) in Spanish (L2) were also considered. The LDT is a task that asks participants to identify real words versus nonwords. A list of 100 real Spanish words and 100 nonwords were created for the LDT. Real words were low frequency but familiar 4-5 letter concrete Spanish nouns. All nonwords were pronounceable, matched orthographic neighbors to the real Spanish words. Since all nonwords adhered to Spanish phonotactic and orthographic rules, accurate performance on the LDT required specific knowledge of Spanish vocabulary, and was therefore a valid indicator of L2 skill. Each participant saw 50 real Spanish words and 50 nonwords (which were the matched pairs of the real words that they did not see). Participants were instructed to use a game pad to press “yes” when the words were real Spanish words, and “no” when the words were not real Spanish words. Two independent measures of cognitive control, the Simon Task and the Operation Span Task, were also used to assess whether participants in the two groups differed on cognitive resources, as well as language experience. These tasks are described in more detail in the Experimental Overview section below.
Table 1. Participant Characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Less Proficient</th>
<th>More Proficient</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>12</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>Study abroad experience</td>
<td>No</td>
<td>Yes</td>
<td>12 no; 12 yes</td>
</tr>
<tr>
<td>Average self rating for English (LHQ)</td>
<td>9.72 (SD = 0.34)</td>
<td>9.93 (SD = 0.13)</td>
<td>9.83 (SD = 0.27)</td>
</tr>
<tr>
<td>Average self rating for Spanish (LHQ)</td>
<td>6.43 (SD = 1.16)</td>
<td>8.08 (SD = 0.92)</td>
<td>7.26 (SD = 1.32)</td>
</tr>
<tr>
<td>Average LDT score*</td>
<td>66%</td>
<td>83%</td>
<td>74%</td>
</tr>
<tr>
<td>Simon Effect*</td>
<td>45.56</td>
<td>38.74</td>
<td>42.30</td>
</tr>
<tr>
<td>Operation span**</td>
<td>43.27</td>
<td>47.64</td>
<td>45.45</td>
</tr>
</tbody>
</table>

* Due to timing constraints, one participant did not complete the LDT or Simon Task, so the overall average and average for the more proficient group are reported for 23 out of 24 and 11 out of 12 subjects, respectively for these measures.

** Due to timing constraints, two participants did not complete the Operation Span Task, so the overall average and average for both groups are reported for 22 out of 24 and 11 out of 12 subjects, respectively for these measures.

The mean values on the proficiency measures and results from the two tests of cognitive control are provided in Table 1 for each group, as well as overall averages. Independent samples t-tests were run to compare the groups based on the variables included in Table 1 above in PASW Statistics 18 (SPSS, Inc.). Results are provided with unequal variances assumed when the Levene's test for unequal variances for that comparison reached significance, although uncorrected degrees of freedom are reported. In addition to differing in their study abroad experience, the two groups differed
significantly on their self-ratings in Spanish, $t(22) = -3.868, p = 0.001$, and in their LDT accuracy, $t(21) = -3.961, p = 0.001$. The more proficient group rated their Spanish skills higher and performed better at discriminating between real Spanish words and non-words than the less proficient group. There was also a trend for the less proficient group to rate their English skills lower than the more proficient group, $t(22) = -2.087, p = 0.055$, although both groups rated their English proficiency highly. The two groups did not differ significantly on the Simon task, $t(21) = 0.667, p = 0.512$, or the operation span task, $t(20) = -1.070, p = 0.298$, suggesting that they were well-matched in terms of cognitive resources.

**Experimental Overview**

All participants completed one experimental session lasting approximately two and a half hours. Informed consent was obtained prior to starting the experiment. Participants also filled out two questionnaires, one of which was the LHQ. The second questionnaire requested demographic information, assessed handedness, and reinforced that the participant met the inclusion criteria for the experiment. After completing the questionnaires, participants then completed four separate computer-based tasks throughout the experiment. Participants first completed the translation recognition task, which was the primary task of interest in the current study. During the translation recognition task a Spanish word was presented on the screen followed by an English word. The stimuli were presented on a computer screen and participants made responses on a game pad. They were told to push the “yes” button on the game pad if the English word was a direct translation of the Spanish word, and the “no” button if the English word was not a direct translation. Task 1 consisted of 4 blocks in which the participant
saw 80 pairs of Spanish and English words in each block. Spanish words were presented on the monitor for 500 ms. They were followed by a blank screen for 250 ms and then the presentation of the English word for 500 ms. Another blank screen was presented for 600 ms followed by a plus sign. The plus sign was centered on the screen for a duration of 2000 ms and followed by another blank screen for 200 ms before the next Spanish word was presented. Participants were instructed to blink only when the plus sign was on the screen in order to reduce artifacts from eye blinks. Figure 1 provides a schematic of a typical trial.

**Figure 1.** Translation recognition task: Time course and organization of stimuli from the onset of one critical item to the next.

The second and third tasks were the LDT and the Simon task (Simon & Ruddell, 1967). These tasks were completed in a sound-attenuated room with dim lighting, and ERP and behavioral data were recorded for all three. The LDT is described in detail.
above. In the Simon task participants saw small red and blue boxes flash in different locations on the computer screen. When the participant saw a red box on the screen, he or she was to respond by pressing the red button on his or her game pad as quickly as possible, and when a blue box was presented, he or she pressed the blue button on the game pad. The response was to correspond to the color of the box, regardless of the location of the box on the screen. The Simon task is considered a measure of inhibitory control, and performance on the Simon task is represented by the “Simon effect”, which can be described as the difference in reaction times (RTs) between trials in which the position of the stimulus was congruent with that of the response compared to trials in which the positions of the stimulus and response were incongruent. Specifically, to calculate the Simon effect RTs for congruent trials are subtracted from RTs for incongruent trials since RTs for incongruent trials are usually slower, larger numbers.

The final task was an operation span task performed in English, which is a measure of working memory capacity (Turner & Eagle, 1989). In this task the participant solved a sequence of mathematical expressions that were each followed by the presentation of an English word on the screen. At the end of the sequence, the participant was asked to recall the English words that were presented. The operation span task was completed in a separate sound-attenuated room, and only behavioral data were recorded. The operation span data reported in Table 1 correspond to the number of English words recalled.

**Stimuli**

In the translation recognition task the stimuli were Spanish-English word pairs that were either correct or incorrect translations. All Spanish and English words were
common nouns. All words had nine or fewer letters. No cognates were used (i.e., words with similar form and meaning across languages, such as problema and problem), and words were not repeated throughout conditions. Each participant saw 320 Spanish words and 320 English words throughout the translation recognition task. As described above, the words were presented one at a time, so that the participant always saw a Spanish word followed by an English word. There were a total of 160 “yes” trials (i.e. correct Spanish to English translations) and 160 “no” trials (i.e. incorrect translations). The “no” trials were comprised of four critical conditions. Each participant saw forty of each critical condition to make up the 160 “no” trials. The yes trials and four critical conditions are represented in Table 2 and the critical conditions are described below.

The first critical condition was semantically related distractors (1). Word pairs were considered semantically related when the English word was related in meaning to the correct translation of the Spanish word presented. Words could be related associatively or categorically for this condition. The second condition was semantic controls (2) for the semantically related distractors. All control words were found using the MRC Psycholinguistic Database (Coltheart, 1987). They were matched to the semantically related distractors based on their number of letters and Kucera and Francis (1967) written frequency (KFFRQ) values. The number of letters between the semantically related words and their controls were kept the same when possible, although a few were one letter shorter or longer than the critical items they were matched to. Control words were chosen if they had KFFRQ values within three digits of the semantically related words with lower frequency values (i.e. between 0 and 10). As the KFFRQ value increased, greater variability was allowed between the KFFRQ values for
semantically related words and their matched control words. For the most part, frequencies were kept within five to ten, with some exceptions. Mean word length and KFFRQ values for semantically related words versus control words can be found in Table 3.

The third critical condition was translation related distractors (3). This condition was comprised of Spanish words followed by English words that were related in form to the correct translation. The majority of translation related words differed in only one or two sounds to the correct translation, but in some cases words were less similar but as closely related as possible (i.e. slavery for the correct translation silver). Translation controls (4) made up the fourth and final critical condition. These word pairs served as controls for the translation related word pairs. Like the semantic control words, they were found using the MRC Psycholinguistic Database and matched to the translation related distractors based on word length and KFFRQ (see Table 3). Separate controls were matched to each critical condition, because it was impossible to perfectly match the lexical characteristics of the English words across semantic and translation conditions, although the same Spanish words appeared across conditions.
Table 2. Sample stimuli for yes trials and critical conditions used in Task 1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stimulus Examples</th>
<th>Correct Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spanish word</td>
<td>English word</td>
</tr>
<tr>
<td>1. Semantically related</td>
<td>PLAYA</td>
<td>SAND</td>
</tr>
<tr>
<td>2. Semantic controls</td>
<td>PLAYA</td>
<td>ROOT</td>
</tr>
<tr>
<td>3. Translation related</td>
<td>PLAYA</td>
<td>BATCH</td>
</tr>
<tr>
<td>4. Translation controls</td>
<td>PLAYA</td>
<td>CANON</td>
</tr>
<tr>
<td>5. Yes trials (fillers)</td>
<td>JUGUETE</td>
<td>TOY</td>
</tr>
</tbody>
</table>
Table 3. *Average word length and KFFRQ for each stimulus list (order) across critical conditions in Task 1.*

<table>
<thead>
<tr>
<th>Order</th>
<th># of letters</th>
<th>KFFRQ</th>
<th># of letters</th>
<th>KFFRQ</th>
<th># of letters</th>
<th>KFFRQ</th>
<th># of letters</th>
<th>KFFRQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.30</td>
<td>86.08</td>
<td>5.30</td>
<td>90.95</td>
<td>4.65</td>
<td>57.00</td>
<td>4.63</td>
<td>60.31</td>
</tr>
<tr>
<td>2</td>
<td>5.10</td>
<td>84.65</td>
<td>5.10</td>
<td>80.18</td>
<td>4.43</td>
<td>63.13</td>
<td>4.43</td>
<td>61.05</td>
</tr>
<tr>
<td>3</td>
<td>5.38</td>
<td>90.21</td>
<td>5.40</td>
<td>90.00</td>
<td>4.53</td>
<td>63.78</td>
<td>4.50</td>
<td>64.21</td>
</tr>
<tr>
<td>4</td>
<td>5.28</td>
<td>82.65</td>
<td>5.28</td>
<td>88.85</td>
<td>4.73</td>
<td>62.24</td>
<td>4.75</td>
<td>62.24</td>
</tr>
</tbody>
</table>

All participants saw the same 320 Spanish words, but were placed in one of four orders that determined which English translations were presented. In order to formulate the four orders, a database was constructed that included 160 correct Spanish to English translations and 160 Spanish words followed by the four critical conditions mentioned. Four randomized and counterbalanced lists were then produced so that based on the order in which a participant was placed, he or she would see the same 160 correct translations, but different critical conditions for a given Spanish word. Table 3 shows that average word length and KFFRQ values for the four critical conditions were similar between the four orders.
**EEG Recording**

The experiment utilized an electrode cap (Electro-Cap International, Eaton, OH) with embedded electrodes to record twenty-nine channels of EEG activity. Eleven electrodes measured from standard international 10-20 system sites. Three of the eleven electrodes were located on the midline at frontal (Fz), central (Cz), and parietal (Pz) sites. The remaining eight 10-20 system locations were as follows: left and right hemisphere frontal (F2/F4), central (C3/C4), parietal (P3/P4), and temporal (T3/T4). Ten additional sites were based on the Modified Combinatorial Nomenclature system. Two were midline sites: the frontal pole (FPz) and the occipital pole (Oz); four were left and right hemisphere fronto-central sites (FC1/FC2, FC5/FC6); and four were left and right centro-parietal sites (CP1/CP2, CP5/CP6). Additionally, eight modified 10-20 system sites were recorded from lateral positions 33% of the distance from FPz to T3/T4 (FP1’/FP2’), 67% of the distance from FPz to T3/T4 (F7’/F8’), 33% of the distance from Oz to T3/T4 (O1’/O2’), and 67% of the distance from Oz to T3/T4 (T5’/T6’). There were also two loose eye electrodes in place to monitor eye movements by measuring the electro-oculogram (EOG). Eye blinks were recorded by an electrode placed just below the left eye (LE), while horizontal eye movements were accounted for by an electrode placed next to the right eye (HE). All electrodes were referenced to a third loose electrode located at the left mastoid (A1). A fourth loose electrode was placed on the right mastoid (A2) to identify any differential mastoid activity. The electrode placement is represented in Figure 2.

Before beginning the experiment, all impedances for scalp and mastoid electrodes were lowered to less than 5 kilo-ohms (kΩ). The eye electrode impedances were lowered
to less than 20 kΩ. (As noted above, one participant was excluded due to an inability to meet these impedance criteria.) An SA amplifier system with a bandpass of 0.1 to 40 hertz (Hz) was used to amplify the EEG recordings, and the EEG was sampled continuously at a rate of 200 Hz throughout the experiment.

![Figure 2. 32-channel electrode montage.](image)

Data Analysis

Behavioral and electrophysiological measures were used to assess performance on the translation recognition task. Only correct responses from the “no” trials were used for analyses. Descriptive and inferential statistics were computed to compare performance between the less proficient and more proficient groups. All analyses were performed using PASW Statistics 18 (SPSS, Inc.). Significant results are reported as those with a $p$ value of less than 0.05, and “trends” are reported with $p$ values between 0.05 and 1.0.

Behavioral data analysis. Separate means for reaction time (RT) and percent accuracies for each of the four critical conditions were computed for each participant. Outliers occurring less than 200 ms or more than 2000 ms post stimulus onset were eliminated from the analysis. RTs slower than 2.5 standard deviations above a given
participant’s mean RT were also considered outliers and thus were excluded from analysis. The magnitude of interference between translation related distractors and their matched controls and between semantically related distractors and their matched controls were obtained by subtracting the values for control conditions from those for the relevant distractor conditions. ANOVAs were also used to compare condition (control versus distractor for each type, semantic or translation) within subjects and to assess proficiency as a between-subjects factor.

**ERP data analysis.** Each participant’s ERP data from all electrode sites for critical conditions were extracted from the continuous EEG. Separate ERPs were averaged for each subject for each critical condition (semantically related, semantic controls, translation related, and translation controls). EEG waveforms began with a 100 ms baseline prior to stimulus onset and were evaluated through 900 ms post-stimulus onset. Again, only correct responses were included in the analyses. The 24 participants included in the final analyses had an average percent accuracy of 82% across all critical conditions. The range for individual subjects for a given critical condition was between 62.9% and 100%. An automated artifact rejection algorithm was initially applied to the data to eliminate trials with eye blinks and other artifacts. When the total number of usable trials for any critical condition was less than 70% of the correct responses, the thresholds for artifact detection were manually adjusted to maximize the useable trials for a given participant, while maintaining appropriate rejection of artifact-containing trials. Any participants whose data still included fewer than 70% useable trials were excluded, as described in the Participants section. Trials with unrecoverable eye blinks and muscle artifacts were eliminated from the averages and were therefore not included in the ERP
results (although these trials did contribute to the behavioral data). Individual ERPs were grand-averaged to assess performance on the task for the group as a whole. Two additional grand averages were made to visualize the effects of proficiency level; one for the less proficient group of participants and another for the more proficient group of participants.

Based on visual inspection of the results and on previous literature, three time windows were selected for analysis of the mean amplitudes: The P200 was analyzed between 150-300 ms; the N400 was analyzed between 300-450 ms; and a late positive component (LPC) was analyzed between 450-700 ms. Separate repeated measures ANOVAs were performed to compare translation related conditions to their controls (translation interference), and semantically related conditions to their controls (semantic interference). Additionally, separate analyses were done for different sets of electrode sites to allow for analysis of the topographic distribution of effects. The following sets of electrode sites were analyzed: midline (FPz, Fz, Cz, Pz, Oz) and three sets of lateral electrode sites (inner circle: FC1/FC2, C3/C4, CP1/CP2; middle circle: F3/F4, FC5/FC6, CP5/CP6, P3/P4; and outer circle: FP1'/FP2’, F7’/F8’, T3/T4, T5’/T6’, O1’/O2’). The factors of interest for each set of analyses were experimental condition (control versus distractor for a given type of item), hemisphere (when appropriate), and electrode site. Additionally, proficiency (i.e., less proficient versus more proficient) was entered into each analysis as a between subjects factor. Only condition or proficiency main effects and interactions between condition or proficiency and the other variables are reported, since main effects and interactions including only electrode site and/or hemisphere are not of interest in the current study. For any comparison with more than two levels, the
Greenhouse-Geisser correction for non-sphericity was applied, although degrees of freedom reported are uncorrected. Post-hoc analyses were performed when interactions found between variables were statistically significant.

Results

Behavioral Data

Behavioral results are presented in Figures 3 and 4. Figure 3 shows the participants’ reaction time (RT) results while Figure 4 shows their percent accuracy. There was a significant difference in RTs between the semantically related distractors and their controls, $F(1, 22) = 29.461, p < 0.000$, indicating that participants were slower to reject the semantically related distractors than their controls. A significant difference was also found in RTs between translation related distractors versus their controls, $F(1, 22) = 16.384, p < 0.001$, indicating that participants were slower to reject the translation related distractors than their controls. There was no main effect of proficiency or interaction between condition and proficiency in the RTs.

There was a significant difference between semantically related distractors and their controls found in the percent accuracy data $F(1, 22) = 140.930, p < 0.000$. This finding indicates that participants were more accurate in rejecting semantic controls than semantically related distractor items. There was a trend toward a main effect of proficiency $F(1, 22) = 3.332, p < 0.082$, which likely reflects the fact that the more proficient participants were generally more accurate than the less proficient participants on these items. There was no significant interaction between proficiency and condition found in the semantic conditions. There also was a significant difference between translation related distractors and their controls $F(1, 22) = 145.809, p < 0.000$, which can
be attributed to the fact that participants were more accurate in rejecting translation controls than translation related distractor items. There was not a significant main effect of proficiency for translation conditions, nor was there a significant interaction between condition and proficiency.

**Figure 3.** Behavioral interference for mean reaction times (in milliseconds) for translation related distractors versus translation controls and semantically related distractors versus semantic controls for the less proficient and more proficient groups. Each bar represents the difference between values for a given distractor condition and the matched control condition.
Figure 4. Behavioral interference for percent accuracy for translation related distractors versus translation controls and semantically related distractors versus semantic controls for the less proficient and more proficient groups. Each bar represents the difference between values for a given distractor condition and the matched control condition.

Visual Inspection of ERP Data

All Participants

Figures 5 and 6 show the ERP waveforms for all participants at all 29 scalp electrode sites. Figure 5 shows the translation related distractors versus translation controls, while Figure 6 shows the semantically related distractors versus semantic controls. All conditions elicited an N100 component (peaking around 110 ms after the onset of the English word), a P200 (peaking around 210 ms), an N400 (peaking between 350 and 400 ms), and a late positive component (LPC) beginning around 500 ms. Translation related trials elicited a slightly larger N400 and a larger LPC compared to translation controls (See Figure 5). In contrast, semantically related trials elicited a smaller N400 in comparison to semantic controls (See Figure 6). Semantically related
trials elicited a slightly larger LPC compared to semantic controls across many frontal and midline sites, but no effect, or a small effect in the opposite direction, was seen at some other sites.

**Figure 5.** Grand average ERPs elicited by translation related distractors and translation controls for all 24 participants. Note that negative is plotted up.

**Figure 6.** Grand average ERPs elicited by semantically related distractors and semantic controls for all 24 participants. Note that negative is plotted up.
Less Proficient Group

Figures 7 and 8 show the ERP waveforms for the less proficient group at all 29 scalp electrode sites. Figure 7 shows the translation related distractors versus translation controls, while Figure 8 shows the semantically related distractors versus semantic controls. Similar results to those described for all participants for the translation conditions were found in the less proficient group (See Figure 7). Again, the translation related distractors and translation controls elicited an N100 (peaking around 125 ms), a P200 (peaking around 110 ms), an N400 (peaking around 400 ms), and a LPC (beginning at 500 ms). Translation related distractors peaked approximately 25 to 50 ms after translation controls at the P200. The N400 was slightly more negative for translation distractors than for translation controls. Additionally, the LPC was more positive for translation related distractors versus translation controls. The following effects for semantic conditions were found in the less proficient group. Both semantically related distractors and semantic controls elicited an N100 (peaking at 110 ms), a P200 (peaking at 210 ms), an N400 (peaking between 375 and 400 ms), and a LPC (beginning around 500 ms). The N100 was slightly more negative for semantic controls compared to semantically related distractors. The N400 was more negative for semantic controls compared to semantically related distractors. Finally, the LPC was more positive for semantically related distractors compared to controls, with the exception of a slightly more negative component found at the lateral temporo-parietal areas of the scalp.
Figure 7. Grand average ERPs elicited by translation related distractors and translation controls for the less proficient group (n=12). Note that negative is plotted up.

Figure 8. Grand average ERPs elicited by semantically related distractors and semantic controls for the less proficient group (n=12). Note that negative is plotted up.

More Proficient Group

Figures 9 and 10 show the ERP waveforms for the more proficient group at all 29 scalp electrode sites. Figure 9 shows the translation related distractors versus translation
controls, while Figure 10 shows the semantically related distractors versus semantic controls. For translation conditions, the more proficient group elicited an N100 (peaking at 110 ms), a P200 (peaking at 225 ms), an N400 (peaking around 400 ms), and an LPC, specifically at the center and right side of the scalp (beginning at around 500 ms).

Translation related distractors were slightly more negative at the N100, slightly more positive at the P200, more negative at the N400 (specifically at the frontal sites), and more positive on the LPC in comparison to translation controls. Additionally, translation related distractors peaked approximately 25 to 50 ms after translation controls on the P200 and N400. Both semantic conditions elicited the following components: N100 (peaking at 100 ms), P200 (peaking at 225 ms), N400 (peaking at 350 ms), and the LPC (beginning at approximately 500 ms). Semantically related distractors were slightly more positive at the N400 than semantic controls. Semantically related distractors showed very subtle effects on the LPC. Semantically related distractors were slightly more positive than semantic controls toward the center of the scalp and on posterior sites, while the semantic controls were slightly more positive at outer electrode sites. The observed pattern of results is characterized with the ANOVAs below.
**Figure 9.** Grand average ERPs elicited by translation related distractors and translation controls for the more proficient group (n=12). Note that negative is plotted up.

**Figure 10.** Grand average ERPs elicited by semantically related distractors and semantic controls for the more proficient group (n=12). Note that negative is plotted up.

Figure 11 compares the translation conditions for the less proficient and more proficient groups at the three midline electrode sites (Fz, Cz, Pz), and Figure 12 compares
the semantic conditions for both groups at midline sites. Both groups showed an enhanced P200 and enhanced N400 for translation related distractors compared to translation controls. Additionally, translation related distractors were more positive on the LPC in comparison to their controls; however, the enhanced positivity was more robust in the less proficient group, specifically at anterior sites. Effects for semantic conditions were more similar between groups, and were characterized by an attenuation of the N400 component for semantically related distractors compared to controls. Additionally, semantically related distractors were more positive than semantic controls at the LPC in both groups at these sites, although these effects are again somewhat more robust for the less proficient group.

**Figure 11.** Grand average ERPs for translation related distractors and translation controls for (A) the less proficient group and (B) the more proficient group at three midline sites, Fz, Cz, and Pz. Note that negative is plotted up.
Figure 12. Grand average ERPs for semantically related distractors and semantic controls for (A) the less proficient group and (B) the more proficient group at three midline sites, Fz, Cz, and Pz. Note that negative is plotted up.

Statistical Analysis of ERP Data

Translation Interference Effect

P200. There was a significant interaction between condition and electrode for the outer circle electrode sites ($F(4, 88) = 3.655, p < 0.029$). This result is consistent with the observation of a slightly larger P200 for translation related distractors than controls. Further paired $t$ tests revealed trends toward significant differences in P200 mean amplitude between translation related conditions and translation related controls at FP1/2 ($t(23) = 2.040, p = 0.053$) and T5/6 ($t(23) = -1.857, p = 0.076$), but no significant results at other sites. There was also a trend toward a significant interaction between condition, hemisphere, and proficiency at the inner circle electrodes ($F(1, 22) = 3.682, p = 0.068$) and at middle circle electrodes ($F(1, 22) = 3.118, p < 0.091$). There were no other significant effects in this analysis window.

N400. There was a significant main effect for proficiency at the inner circle electrode sites ($F(1, 22) = 4.497, p = 0.045$), and a trend toward significance at midline
electrode sites \((F(1, 22) = 3.967, p = 0.059)\). This result reflects the fact that there were differences in the effects seen on the N400 epoch between the less proficient and more proficient groups. Specifically, the more proficient group showed an enhanced negativity for translation distractors versus translation controls at these sites in comparison to the less proficient group where the effect was not as prominent.

There was a significant interaction between condition and electrodes at the inner circle electrode sites \((F(2, 44) = 7.856, p = 0.003)\) and middle circle electrode sites \(F(3, 66) = 5.511, p = 0.014\), and a trend toward significance at the midline electrode sites \((F(4, 88) = 2.983, p = 0.056)\). These results likely reflected the fact that differences between conditions (i.e., a larger N400 for the translation related distractors than controls) were more pronounced towards the front and center of the scalp. Further paired \(t\) tests showed no significant differences between the conditions at inner circle electrodes \((t(23) < 1)\), and over the middle circle electrodes, significant differences in N400 mean amplitude were found only at FC5/6 \((t(23) = -2.515, p = 0.019)\).

There was a significant interaction between condition and hemisphere at the outer circle electrode sites \((F(1, 22) = 5.604, p < 0.027)\), and a trend toward significance at the inner circle electrode sites \((F(1, 22) = 3.580, p < 0.072)\) and middle circle electrode sites \((F(1, 22) = 3.635, p < 0.070)\). There was also a significant interaction between condition, hemisphere, and electrodes at outer circle electrode sites \((F(4, 88) = 3.516, p < 0.010)\), and a trend toward significance at middle circle electrode sites \((F(3, 66) = 2.697, p < 0.092)\). Follow-up statistics using a simple effects analysis to break down the significant three-way interaction of condition, hemisphere, and electrode at outer circle sites revealed no significant interaction of condition and electrode at either left or right
hemisphere sites. However, in this analysis there was a trend towards a significant effect of condition at left hemisphere sites \((F(1, 23) = 3.267, p = 0.084)\), while there was no effect of condition at right hemisphere sites. This result likely underlies the significant interaction between condition and hemisphere at the outer circle sites. These results are consistent with the observation that the N400 enhancement for the translation related distractors was more pronounced at the fronto-central, left hemisphere outer circle sites.

Finally, there was a significant interaction between condition, hemisphere, and proficiency at middle circle electrode sites \((F(1, 22) = 5.305, p = 0.031)\) and outer circle electrode sites \((F(1, 22) = 4.496, p = 0.045)\). Simple effects analyses were performed for both of these interactions, where condition by hemisphere effects were evaluated separately for each group. At the middle circle sites, no significant condition by hemisphere interaction was observed for the lower proficiency group. However, the high proficiency group did display a significant condition by hemisphere interaction at these sites \((F(1, 11) = 13.519, p = 0.004)\). At the outer circle sites, a similar pattern was noted, with no significant interaction of condition by hemisphere for the lower proficiency group, while this interaction was significant for the high proficiency group \((F(1, 11) = 16.456, p = 0.002)\). In general it appeared that the N400 effects of condition were greater on the left hemisphere for the more proficient participants, while no clear hemispheric asymmetry was present for the less proficient participants.

**LPC.** There was a significant main effect of condition on all electrode sites (midline: \(F(1, 22) = 18.965, p = 0.000\)); inner circle: \(F(1, 22) = 13.065, p = 0.002\)); middle circle: \(F(1, 22) = 8.780, p = 0.007\)); outer circle \(F(1, 22) = 6.030, p = 0.022\). This result reflects the fact that the mean amplitude for the translation related distractors was
more positive than the translation controls on the LPC. There was also a significant main
effect for proficiency at midline electrodes ($F(1, 22) = 5.542, p = 0.028$) and middle
circle electrodes ($F(1, 22) = 4.591, p = 0.043$), and a trend toward significance at inner
circle electrodes ($F(1, 22) = 4.078, p = 0.056$) and outer circle electrodes ($F(1, 22) =
4.031, p = 0.057$). These results show that there were differences in the effects seen
across the LPC between the less proficient and more proficient groups. Specifically, the
average amplitude for the LPC for the less proficient group (across both distractor and
control items) was less positive than for the more proficient group.

There was a significant interaction between condition and electrode at inner circle
electrodes ($F(2, 44) = 5.109, p = 0.022$) and middle circle electrodes ($F(3, 66) = 4.285, p
= 0.040$), and a trend toward significance at outer circle electrodes ($F(4, 88) = 2.786, p =
0.088$). Further paired $t$ tests showed significant differences between conditions for the
LPC at all inner circle electrode sites: FC1/2 ($t(23) = 4.486, p = 0.000$); C3/4 ($t(23) =
3.037, p = 0.006$); and CP1/2 ($t(23) = 2.881, p = 0.008$), although the magnitude of the
difference was somewhat larger towards at the fronto-central sites (FC1/2). Significant
differences were also found across middle circle electrode sites at F3/4 ($t(23) = 4.140, p =
0.000$), FC5/6 ($t(23) = 2.258, p = 0.034$), P3/4 ($t(23) = 2.672, p = 0.014$), and a trend
toward significance was revealed at CP5/6 ($t(23) = 1.789, p = 0.087$).

Finally, there was a trend toward a significant interaction between condition,
hemisphere, and proficiency at middle circle electrodes ($F(1, 22) = 3.383, p = 0.079$).

*Semantic Interference Effect*

*P200.* Although no effects reached statistical significance for this analysis, there
was a trend toward a significant main effect for proficiency at the outer circle electrode
sites ($F(1, 22) = 3.345, p = 0.081$). There was also a trend toward a significant interaction between condition, electrode, and proficiency for the outer electrode sites ($F(4, 88) = 2.597, p = 0.086$).

*N400.* There was a significant main effect for condition at all electrode sites (midline: $F(1, 22) = 8.426, p = 0.008$); inner circle: $F(1, 22) = 9.128, p = 0.006$); middle circle: $F(1, 22) = 7.801, p = 0.011$); outer circle $F(1, 22) = 6.188, p = 0.021$). This result reflects the fact that the semantic controls were more negative than the semantically related distractors on the N400 epoch. There was also a significant main effect for proficiency at the midline ($F(1, 22) = 6.857, p = 0.016$) and at the inner circle electrodes ($F(1, 22) = 5.907, p = 0.024$), and a trend toward significance at middle circle electrodes ($F(1, 22) = 4.238, p = 0.052$). These results show that there were differences in effects for the mean amplitudes on the N400 between the less proficient and more proficient groups. Overall, mean amplitudes for semantic conditions were more negative for the less proficient group compared to the more proficient group, although the relative differences between conditions were similar for each group.

There were trends toward a significant interactions between condition and electrode at the midline ($F(4, 88) = 2.660, p = 0.091$) and inner circle electrode sites ($F(2, 44) = 3.774, p = 0.058$).

*LPC.* There was a significant main effect for proficiency at the midline electrodes ($F(1, 22) = 9.026, p = 0.007$) and the middle circle electrodes ($F(1, 22) = 4.600, p = 0.043$), and a trend toward significance at inner circle electrodes ($F(1, 22) = 3.473, p = 0.076$) and outer circle electrodes ($F(1, 22) = 3.598, p = 0.071$). This difference reflects the tendency for the more proficient group to have more positive waveforms in the LPC
epoch overall. There was also a trend toward a significant main effect of condition at the midline \(F(1, 22) = 3.130, p = 0.091\).

Discussion

The purpose of the current study was to assess how lexical access changes as a result of proficiency level by examining the performance of two groups of English-Spanish bilinguals in a task that required them to determine whether two words were translations across their two languages. Both ERP and behavioral responses were recorded while the participants performed a translation recognition task in which they decided whether an English target word was the correct translation of a previously presented Spanish word. Equal numbers of correct and incorrect translation trials were presented, but critical conditions were all trials in which the English word was an incorrect translation of the Spanish word. Critical conditions were as follows: words that were related in lexical form to the correct translation (“translation related distractors”), words that were related in meaning to the correct translation (“semantically related distractors”), and controls for each distractor type (“translation controls” and “semantic controls”).

Behavioral measures of reaction time and accuracy revealed interference for both distractor types, indicating that the related trials were more difficult to reject as incorrect translations than the unrelated matched controls. According to both reaction time and percent accuracy data, the less proficient and more proficient groups both experienced a larger magnitude of interference for semantically related distractors than they did for translation related distractors. Significant differences between groups were not observed in the reaction time data; the percent accuracy data only showed a trend for the less
proficient group to be less accurate across both types of semantic conditions than the more proficient group. This result is inconsistent with previous behavioral results by Talamas et al. (1999) who found greater interference for translation related distractors in less proficient participants and greater semantic interference in more proficient participants. Our results do not suggest the developmental shift from a reliance on form to a reliance on meaning that was reported in the Talamas et al. study. Instead, our behavioral results suggest that both lexical and conceptual information are active regardless of a bilingual’s level of proficiency. The finding in the present study that conceptual information was active in the less proficient group is consistent with the studies by Altarriba and Mathis (1997), Dufour and Kroll (1995), and Sunderman and Kroll (2006), all of which provided evidence for a direct access to concepts in early L2 learning. Additionally, other studies have found continued access to translation equivalents in skilled bilinguals (Altarriba & Mathis, 1997; Guasch et al., 2008; Guo et al., submitted; Thierry and Wu, 2007). Therefore, it is not surprising that the more proficient group in the present study experienced behavioral interference for translation distractors.

Overall, the ERP results were consistent with the behavioral results in that effects for translation related distractors and semantically related distractors were observed in the ERP waveforms for both groups, confirming that lexical and conceptual information is active regardless of the bilinguals’ proficiency level. However, the ERP results showed different patterns of effects based on proficiency level. For translation conditions, the more proficient group showed an enhanced negativity on the N400 at frontal sites for translation related distractors that was not observed in the less proficient group. There
was also an enhanced late positivity for the translation conditions in both groups, although the magnitude of this difference was larger for the less proficient group, particularly at frontal sites. Semantic conditions elicited results that were more similar between groups, with the most noticeable effect being an attenuation of the N400 component for the semantic distractors, as compared to controls.

In this study, the N400 was shown to be sensitive to effects of both types of distractors, although the distractor type impacted the direction of the effects. The fact that the semantically related distractors elicited an attenuated N400 compared to semantic controls in the current study is consistent with previous ERP findings. These results may be explained by the fact that the participants were aware of the similarities in meaning between the semantically related distractors that were presented and the correct translations. The attenuation of the N400 for the semantic conditions is similar to effects that have been interpreted as facilitation in other tasks, such as semantic priming. However, in the translation recognition task, this apparent facilitation on the N400 was associated with an interference effect in the behavioral data, likely due to the specific demands of the task. In contrast to the ERP results for the semantic distractors, there was an enhanced N400 for translation related distractors compared to translation controls. This result may reflect activation of an inhibitory process when encountering a word that is related in form to the expected target. The frontal distribution of this enhanced negativity and the fact that it was more pronounced in the more proficient group of participants support the notion that participants may have experienced interference from being forced to process a lexical alternative to the correct L1 translation. This effect might suggest that participants had already inhibited these competitors once they
identified the meaning of the L2 word. If this is the case, the source of the observed behavioral interference for semantic versus translation related distractors may differ significantly, despite the similar pattern to the behavioral effects.

In addition to effects on the N400, critical conditions elicited effects that were evident later in the decision making stage. In the current study, there was an enhanced LPC for distractor items in the translation conditions for both groups. For the most part, semantic distractors were also more positive than semantic controls at the LPC for both groups. The enhanced positivity for distractor items on the LPC may reflect the participants’ process of double-checking their responses when items were related in form or meaning to the correct translation equivalents. The argument that participants were “rechecking” responses was also made in the study by Guo et al. (submitted), which also found late effects for both types of conditions. However, the Guo et al. study found marked differences in the direction of effects on the LPC for semantic conditions based on scalp position, while only subtle topographic differences were observed in the current study.

The ERP effects observed in the current experiment provide evidence that both lexical and conceptual information were active in the translation recognition task, regardless of the bilingual’s proficiency level. However, the present experiment shows that the pattern and time course for lexical and semantic interference effects vary as a function of proficiency level, with the translation form interference showing the greatest differences between the two groups. In the less proficient bilinguals, translation form interference was most evident later in the decision making stage as evidenced by the late positivity, especially at frontal sites. Translation interference was also present on the LPC
in the more proficient group. However, the translation effect in the more proficient group was more prominent earlier in the waveform, and was characterized by an enhanced N400 at frontal sites. The fact that translation interference was present in the ERPs for the more proficient group is inconsistent with some of the previous behavioral studies that reported translation interference only in less proficient participants (Sunderman & Kroll, 2006; Talamas et al., 1999). However, there are behavioral studies that found evidence for continued access to the L1 translation equivalent in proficient bilinguals (e.g., Altarriba & Mathis, 1997; Guasch, et al., 2008), as well as recent ERP studies (Thierry and Wu, 2007; Guo et al., submitted). These discrepancies in findings between studies may be explained by differences in experimental paradigms. One such difference is that of the stimulus onset asynchrony (SOA) (i.e., the time interval from onset of the first word to onset of the second word). Studies that did not find activation of the L1 translation equivalent in highly skilled bilinguals used a shorter SOA (e.g., 500 ms in Sunderman & Kroll, 2006 and Talamas et al., 1999), while studies that did find activation of the L1 translation equivalent used a longer SOA (e.g., 750 ms in the Guasch et al., 2008, Guo et al., submitted, and in the current study; 1100 ms in Thierry & Wu, 2007). It is possible that the increased SOA encouraged these highly proficient bilinguals to activate the translation equivalent after they accessed the concept for the L2 word. As previously mentioned, this interpretation is consistent with results from the study by Guo et al. (submitted), where they directly compared the effects of a longer SOA (750 ms) to a shorter SOA (300 ms). As was expected, behavioral data showed both form related and semantic interference in participants who performed the task with the longer SOA, but only semantic interference was found in the RTs for the participants utilizing the shorter
SOA. Guo et al. concluded that while skilled bilinguals have the ability to directly access the meaning of L2 words, the translation equivalent may be activated as a consequence of an interactive lexical system. Therefore, the presence of translation interference for the more proficient group in the current study is not surprising, as the increased SOA likely encouraged access to the translation equivalent after activation of the conceptual information had been acquired.

An advantage of the current study is that it went beyond the Guo et al. (submitted) study in evaluating the performance of a lower-proficiency group. This feature allows for more direct evaluation of the developmental claims of the RH M in the current study. Thus, in addition to providing evidence for continued access to the translation equivalent in highly skilled bilinguals, the current study provided evidence that translation equivalents may be activated for different reasons for participants of different levels of proficiency. Another important finding in the current study was ERP evidence for access to conceptual information in the less proficient group. This result is consistent with several behavioral studies that found evidence for activation of conceptual information in bilinguals at the earlier stages of L2 learning (e.g., Altarriba and Mathis, 1997; Dufour and Kroll, 1995; and Sunderman and Kroll, 2006).

Taken together, the behavioral and ERP data from this experiment have several important implications. First, they provide evidence that lexical and conceptual information are active regardless of a bilingual’s proficiency level. Although the experiment proved that both groups of bilinguals had access to lexical and conceptual information, the ERP data showed that the way in which they utilized this information varied as a function of L2 skill. Specifically, the time course of activation differed based
on proficiency level, especially for the translation conditions. Results suggest that the locus of interference for translation related distractors may shift over development. It is hypothesized that the less proficient learners may have considered the translation distractors as potential translations, leading to greater effects in the decision making stage for these items. In contrast, the more proficient learners may have found it difficult to process them when they were presented because they had already inhibited these competitors. The enhanced frontal N400 for translation distractors in the more proficient group is consistent with an inhibitory process affecting lexical competitors of the translation. In regard to the semantic information, effects on the N400 in both groups revealed that learners, not just skilled bilinguals, are sensitive to conceptual information provided by L2 words and that the processes underlying this conceptual access may be fundamentally similar in both less and more proficient bilinguals.
References


