Examining the Bilingual Mental Lexicon Through Associative Priming

Matthew Perham Rollins, BA

Submitted in partial fulfilment of the requirements for the degree of

Master of Arts, Psychology

Faculty of Social Sciences,
BROCK UNIVERSITY
St. Catharines, Ontario

Abstract

Research examining the associations between words in the monolingual versus bilingual mind has employed various models to examine differences in lexical organization, with varying degrees of success. The paradigms used have primarily been word association and semantic priming with a Lexical Decision Task (LDT). This thesis research has focused on the latter method, with an online data collection method using *Testable*. One distinction of this thesis research has been the types of semantic associations used for priming, namely syntagmatic and paradigmatic associations, which refer to either word context in a sentence, or word categories respectively. The control condition used from which facilitation effects were calculated was unrelated primes. In addition, a phonetic (or "clang") priming condition was included as it was felt that it might tap into an important aspect of lexical organization for those who have English as a second language (L2). Recruitment was for native English-speaking monolinguals, native English-speaking bilinguals (who also speak a variety of other languages), and non-native English-speaking bilinguals (also from a range of language backgrounds) to participate. Results indicated that the paradigm was successful in gathering information about lexical associations in all three language groups. There was significant semantic facilitation across all language groups for both syntagmatic and paradigmatic associative primes, with these effects not differing from each other. Interestingly, only the L2 group showed significant facilitation from clang primes. Overall, the absolute priming effect was smaller than anticipated, despite reaching statistical reliability, suggesting possibilities to refine the display times of primes or targets. Other hypotheses concerned potential effects of participants' context for L2 language learning and also attempts to address the main research question with the use of a classic word association task;

however, both fell victim to the vagaries of online data collection. Nevertheless, the method and the software provide some hope for continued research in some aspects of the monolingual versus bilingual mental lexicon.

Acknowledgements

I'd like to thank my co-supervisors, Dr. Sid Segalowitz and Dr. Gary Libben, for all their efforts to guide me through this thesis project. Their instruction and counsel gave me the perspective and resolve to whole-heartedly explore this field of study. Additionally, I thank my committee member Dr. Karen Arnell for her input and suggestions, reminding me to look at my research from multiple perspectives. I also thank the countless lab members in the *Words in the World* Lab, and others who I've worked alongside and learned from both implicitly and explicitly. Lastly, I'd like to thank my mother, Maria Evelyn Jovel-Rollins, whose emotional and intellectual support I relied on countless times on my academic journey.

I stand on the shoulders of giants.

TABLE OF CONTENTS

Abstract…… 2

Acknowledgements ……. 4

List of Tables…… 7

List of Figures…….. 8

List of Appendices ……... 9

Introduction…….. 10

Monolingual Mental Lexicon …… 11

Bilingual Mental Lexicon …….. 13

Syntagmatic – Paradigmatic Shift …….. 15

Individual and Demographic Factors ……. 16

Gaps in the Literature ……. 18

Current Study …….. 20

Rationale …… 23

Methods ……... 26

Participants ……. 26

Stimuli …… 27

Procedure ……... 32

Analyses ……. 34

Results ……. 37

Pilot Study …….. 37

Main Study ……. 37

Native Language …….. 40

Participant Language Status – L1 Monolinguals vs. L1 Bilinguals ……. 42

Participant Language Status – L1 Bilinguals vs L2 Bilinguals ……... 43
Demographic Factors ……..... 45
Word Association Task …… 46

Discussion …….. 47

Mental Lexicon Formation …… 48

Small Priming Effect Problem ……..... 49

Motivation & Gamification …….. 51

Limitations ……. 52

Unexpected Results and Issues to Explore …… 53

Conclusion …….. 55

References…… 56

List of Appendices ……... 60

List of Tables

Table 1 – *Sample of Target words and their associated primes* ……. 28

Table 2 – *Sample of Target non-words and their associated primes* …….. 29

Table 3 - *Sample of Counterbalanced Prime Lists for Participants*…… 30

Table 4 – *Mean Response Times for Lexical Decision Task by Prime Type and Participant Language Status.* …….. 38

Table 5 – *Facilitation Effects for Lexical Decision Task by Participant Language Status*. 39

Table 6 – *Complete List of Target Words* ……... 60

Table 7 - *Complete List of Target Non-Words*……. 64

Table 8 – Participants' Accuracy for Target Words in the Pilot Study …….. 72

Table 9 – Participants' Accuracy for Target Non-Words in the Pilot Study …….. 73

Table 10 - *Estimated Marginal Means Contrasts between Linguistic Groups* ……... 74

Table 11 - *Estimated Marginal Means between Prime Types' Facilitation Effects* …….. 74

Table 12 *LMER Fixed Effects on Mean Reaction Time between Participants' Native Language* ……. 75

Table 13 – *LMER Fixed Effects on Mean Reaction Time according to Participants' Linguistic Status* …… 76

Table 14 – *LMER Fixed Effects of Reaction Times in L2 Bilinguals* ……. 77

Table 15 – *LMER Fixed Effects of Reaction Times in L1 Bilinguals* ……. 77

Table 16 – *LMER Fixed Effects of Reaction Times in L1 Monolinguals* …… 78

Table 17 – *LMER Fixed Effects of Reaction Times in All L1 Speakers* …….. 78

List of Figures

Figure 1 – *Hypothesis 1. Facilitation Effects of Associative Primes across Native Language.* …….. 22

Figure 2 – *Hypothesis 2. Degree of Facilitation of Associative Primes in L2 Speakers across Proficiency levels.* …….. 22

Figure 3 – *Hypothesis 3. Degree of Facilitation of Associative Primes in L2 Speakers across Learning Methods.* ……. 23

Figure 4 – *Sample of Linguistic Profile Questions* …… 31

Figure 5 – *Sample of English Proficiency Questions* …… 32

Figure 6 – *Presentation of Stimuli in the Lexical Decision Task (LDT)* ……... 33

Figure 7 - *Degree of Facilitation between Participant Language Categories across Priming Condition* ……... 40

Figure 8 – *Effects of Prime Type on Reaction Time measured by Native Language* …… 42

Figure 9 - *Effects of Prime Type on Reaction Time measured by Participant Language Status.* ……. 44

List of Appendices

Appendix A- Lexical Decision Task and Word Association Stimuli 60

Appendix B – Language Questionnaires 68

Appendix C – Results Tables 72

Introduction

Knowledge and comprehension of words are core features of language ability. This word knowledge and understanding is fundamentally based in relational connections. Words receive meaning by virtue of their connections to physical objects, abstract concepts, environmental conditions and, perhaps most importantly, other words. It is this last category of inter-word connections that will be the focus of this thesis.

By examining word-to-word connections, we can gain insight into the organization and storage of words in the mind, a concept known as the "mental lexicon". Specifically, the mental lexicon refers to the connected network of words that an individual stores in their mind (Libben, 2021). This construct was conceived in order to explain the mental organization of language, part of an ongoing discussion about language and mind since the 18th century (Odlin, 2005; Pavlenko, 2014). As has been noted by Vaid and Meuter (2017), the fact that most people in the world speak more than one language presents special challenges for theories and models of the organization of the mental lexicon. Until relatively recently, the issue of the bilingual lexicon was seen as a special case, requiring special models that are restricted to bilingualism, and research about the mental lexicon has strongly favoured monolingual models, often at the expense of bilingual ones (Pavlenko, 2014; Vaid & Meuter, 2017). However, Vaid and Meuter (2017) suggest that it is perhaps the monolingual lexicon that should be seen as a special case, because indeed it is the one that is less common. However, this bias has begun to shift, as more research is done on bilingual models and in bilingual settings (Doczi, 2020; Pan & Jarred, 2020; Pavlenko, 2009; Sabourin, Brien, & Burkholder, 2014).

At present, there are several ongoing questions in bilingual lexical research, including how bilinguals' lexical organization for their second language changes over time as they acquire

it. For example, different types of word associations have been employed to distinguish these differences. These include position-based or syntagmatic associations (e.g., “apple pie” associates the two words in a syntactic pairing) and categorical or paradigmatic associations (e.g., “apple” and “pear” are associated through their membership in a semantic category) (Chiu & Lu, 2015). These connections have primarily been studied using word association procedures, either in terms of free recall or word associates format tests where individuals are given a word and asked what other words this brings to mind (Zhang & Koda, 2017). This thesis examines the relative extent of bilinguals’ syntagmatic and paradigmatic organization of lexical components using semantic priming for the reasons indicated below, with the goal of better understanding the interaction between lexical storage, language acquisition, and individual differences.

Monolingual Mental Lexicon

Discussions about language and the brain emerged from 19th century scientists, such as Broca and Wernicke, who sought to identify brain locations associated with language storage and processing models, but without detailed discussion of the network structure of words (Libben, 2021). Despite this early start, serious scientific investigation proposing the concept of a “mental lexicon” only began to proliferate in the 1960s (Collins & Quillian, 1969; 1970), developing models such as the *hierarchical network model* (Collins & Quillian, 1969) and the *semantic feature model* (Smith, Shoben, & Rips, 1974).

These early models focused heavily on categorization of concepts, and how words are associated with concepts. For example, in the *hierarchical network model*, the word ‘bird’ is situated in the superordinate category of ‘animal’ and contains subordinate categories such as ‘pigeon’ and ‘robin’. In this model, connections between adjacent categories take less time to activate than connections between categories further apart in the hierarchy. The *semantic feature*

model, by contrast, views concepts as a series of interconnected features that differ according to a concept's associated features. For example, the concept of a "bird" would have several typically associated features, such as flight, oviparity (egg laying), and feathers. Some species labels in this category would activate more quickly due to higher overlap of features, such as 'robin' or 'pigeon', than other less 'typical' birds, such as 'ostrich' or 'penguin'.

Despite having extensive organization of lexical connections, these models were unable to predict behavioural response times when presented to research participants and relied too heavily on rigid structures (Doczi, 2020). Thus, in response, the *spreading activation model* (Collins & Quillian, 1969, 1970) was also proposed, becoming the most prominent of the monolingual lexical theories. It states that relationships between concepts are dictated by their degree of association, and that the storage was split between lexical information and more mechanical components, such as orthography. Thus, it bears similarities to the *semantic feature model,* but lacks its strict hierarchy, and can also explain the effectiveness of lexical priming behaviours.

In each of the mental lexicon models discussed, a common theme emerges: the primary subject has a monolingual brain. This 'monolingual lens' really only applies to monolinguals (Vaid & Meuter, 2017), though unfortunately it has often affected work examining research participants who speak more than one language. Some have worked around this limitation, by comparing languages in separate monolingual speakers, such as the case of Hopi and English by Whorf (Lucy, 2005) or English and Yucatec (Odlin, 2005), but this still does not address the experience of bilinguals. This lack of bilinguals in these streams of research means that any attempt to apply monolingual models onto bilingual speakers are only conjectures. Bilinguals are

not two monolinguals sharing a brain; thus, other models must be examined in order to understand language organization in the "bilingual brain" (Grosjean, 1989).

Bilingual Mental Lexicon

Research on the bilingual lexicon has lent more support to theories positing a shared lexicon with interrelated concepts (Doczi, 2020; Pavlenko, 2009). However, as mentioned, there is still ongoing debate about the language dynamics of proposed shared systems. For example, early models such as the *word association model* state that words in a speaker's second language (L2) only have direct connections to words in their first language (L1), with no direct association with the underlying concepts. However, this approach does not account for linguistic phenomena such as *Conceptual Transfer.* This occurs in L2 speakers when a concept from their L1 has no equivalent in their L2, and the L1 concept is employed when speaking L2, resulting in L2 errors (Marian & Kaushanskaya, 2007; Odlin, 2005). Research by Malt and colleagues (2015) tested late Mandarin-English bilinguals to see how increased exposure to a second language affected naming patterns, finding that L2 exposure changed both L1 and L2 patterns via *Conceptual Transfer*. Thus, contrary to prior research about the static L1, even L1-concept connections can be altered, leaving this model insufficient. Thus, alternatives such as the *concept mediation model*, where L1 and L2 link to concepts but not to each other, and the *mixed model*, a combination of *word association* and *concept mediation* models, were formulated to try to describe this relationship (Doczi, 2020). These models represent early steps in the study of the bilingual lexicon and its organization, but still cannot fully explain all the potential connections. This nuance is better addressed to some extent by the *Revised Hierarchical model*.

The *Revised Hierarchical model* described the organization as one involving connections between concept and L1, concept and L2, as well as between languages (Kroll & Stewart, 1994). The strength of these links can vary, often with L1-concept connections activating more quickly than L2-concept connections. The model was formulated in order to better explain the connections observed in later L2 learners and highlights the dynamic nature of the lexicon (Doczi, 2020; Sabourin et al., 2014). Here, an L2 may require a stronger connection between L1-L2 in newer learners for scaffolding, but as proficiency increases, so too do the L2-concept connections. Furthermore, by keeping separate semantic and lexical stores, it potentially argues for both separate and combined lexicons based on linguistic level (Sabourin et al., 2014).

However, the mental lexicon itself is a multi-faceted construct, which covers an individual's storage and organization of all their lexical knowledge. Therefore, using the *Revised Hierarchical Model* as a foundation, some researchers have chosen to study some of the bilingual lexicon's smaller components. Simon, Sjerps, and Fikkert (2014) examined the role of phonological components on the organization of the bilingual lexicon, as did Zareva (2007). Other researchers have looked at translation priming, and how grammatical organization in L1 affects the L2 (Sabourin, et al., 2014; Wen & Van Heuven, 2017).

As seen by the range of models, and their constant modification, the evolution of bilingual lexicon models continues as more information is uncovered. Currently, some are proposing the need to develop new paradigms (Libben, 2021), yet the ongoing research of a dynamic mental lexicon in both monolinguals and bilinguals indicates that there are still possibilities to uncover new information about bilinguals and their lexical organization via this paradigm.

Syntagmatic-Paradigmatic Shift

One of the ongoing discussions about mental lexicon development is the similarity between the development of the L2 lexicon compared to the lexicon of a speaker's first language (L1). For example, one phenomenon that occurs in L1 language acquisition is called the "Syntagmatic – Paradigmatic shift". This phenomenon describes the classification of words according to their grammatical class (Nelson, 1977). According to Ervin (1961), children between ages 6 and 8 switch from making continuous syntactic associations (e.g., *apple* $\rightarrow$ *pie* or *apple* $\rightarrow$ *red*) to making categorical associations (e.g. *apple* $\rightarrow$ *pear* or *apple* $\rightarrow$ *fruit*). Nelson (1977) further supports these findings, arguing that this shift results from both a reorganization of the lexicon as well as a change in children's interpretation of the task.

Despite the evidence of the "Syntagmatic – Paradigmatic shift" occurring in L1 learners, the evidence for its occurrence in L2 learners is inconclusive. For example, Zareva (2007) argues that this phenomenon does not occur in adolescent and adult sequential bilinguals, while Clark's (2012) research found that a general trend from syntagmatic to paradigmatic associations does occur in some groups of older bilinguals. However, Clark (2012) also indicates that the bilinguals who exhibited the Syntagmatic-Paradigmatic shift had more frequent and long-term exposure to their L2 than their counterparts. Horiba (2012) suggest that associative preference is language specific, and the learners in some languages will experience this. This lack of consensus suggests that more research may help us identify how syntagmatic and paradigmatic associations form in bilinguals' second language.

Individual and Demographic Factors

When examining the development of the mental lexicon, there are both language-specific features, such as the "syntagmatic – paradigmatic shift", and human factors, such as language proficiency, that are important factors (Ervin, 1961; Nelson, 1977).

Unlike monolingual speakers, bilingual speakers are not a homogenous group, with many individual differences shaping their mental lexicons. For example, Sabourin, Brien, and Burkholder (2014) examined how the age at which a speaker learned their L2 affects the strength of grammatical and translational lexical connections in priming tasks. Here, they examined four different groups of French L2 speakers, from those who learned it alongside their L1 (simultaneous bilinguals) to those who had only begun learning the language. They found that participants who began learning their L2 early in life, either as simultaneous or early sequential bilinguals, would respond to L2-to-L1 translation priming. However, those who were exposed to their L2 as adolescents or adults, would not have the same reaction. These findings indicate that an individual's age of L2 acquisition (AoA) plays a role in the organization and operation of the bilingual mental lexicon.

Another factor found to shape lexical organization is a speaker's L2 proficiency. Zareva (2007) examined the effects of proficiency on quantitative components of the mental lexicon, such as associative strength of lexical connections and total responses, as well as qualitative components, such as paradigmatic and syntagmatic relationships. In this study, the researcher separated the participants into three groups according to their L2 proficiency. The findings indicate that responses and accuracy of quantitative components increased alongside proficiency, while no significant differences were found for qualitative components. Zareva thus argued that proficiency affects the quantity of lexical connections for a speaker, since vocabulary knowledge

increases with proficiency, while having no modulating effect on qualitative components. This latter finding led the researchers to conclude that the qualitative aspects of the lexicon were no different from the lexicons of native speakers.

While some research has broadly focused on L2 speakers' proficiency, others have sought to look at subgroups within this factor. Lee, Kang, and Choi (2018) focused on how the changes in the operation of semantic priming operates in unbalanced bilinguals. This refers to bilinguals with high proficiency in one of their languages, and low proficiency in their other one. Lee and colleagues' research tested Korean-English unbalanced bilinguals on their ability to detect a semantic prime and found that the semantic priming could not occur when the inter-stimulus interval (ISI) between prime and stimulus was too short (10 ms rather than 100ms). While the age of acquisition and proficiency do have overlap, they do measure subtly different attributes of an L2 speakers' linguistic abilities. These two factors have been examined extensively in the development of individuals' mental lexicons (e.g., Borodkin et al., 2021).

A lesser studied factor is the channel or method of L2 Acquisition. This can refer to the medium in which words are first encountered, such as utterances or print (Auer & Bernstein, 2008; Mohsen, 2016), or the learning strategies used to learn the second language (Barcroft, 2008). Research about this topic has only begun to emerge recently, with researchers such as Mohsen (2016) examining L2 acquisition using technological supports, such as annotations, and captions for videos. Some researchers, such as van Hell and Tanner (2012), argue that for a mental lexicon study to be truly applicable to bilinguals, it needs to account for participants' L2 learning background.

Gaps in the Literature

Most of the current research examining the nature of syntagmatic and paradigmatic associations has used word association, a method involving a stimulus word, followed by space for participants to write down an associate (Clark, 2012). While word association has been the prevailing paradigm, semantic priming is an understudied avenue for examining this phenomenon. Semantic priming is '*the improvement in speed or accuracy to respond to a stimulus…when…preceded by a semantically related stimulus (e.g., cat-dog) relative to…a semantically unrelated stimulus (e.g., table-dog).'* (McNamara, 2005, p. 3).

In research, this is usually achieved by pairing a context or experience with a subsequent stimulus, with a follow-up task measuring the behavioural, or neurological effects caused by the prime. One of the most common tasks used in bilingual semantic priming research is the lexical decision task (Neely et al., 1989; Struck & Jiang, 2021). Here, participants are asked to determine if a string of letters (e.g., *house* or *husa*) is a word or not in the target language. Some benefits of measuring this process include its ubiquity in studies of semantic priming portions of speech and language (McNamara, 2005). Semantic priming has previously been implemented in various domains of language, such as linguistic comprehension, lexical recognition, and even representations (e.g., Kealtey & de Gelder, 1992; Lee, Kang, & Choi, 2018; Sabourin et al., 2014; Singh, 2014).

In bilingualism research, semantic priming has been used to measure reaction times and accuracy in various tasks. As mentioned previously, several semantic priming studies have used lexical decision tasks, to measure bilinguals' cognition (de Groot, Borgwaldt, Bos, & van den Eijnden, 2002; Fitzpatrick & Izaura, 2011). Other components of the lexicon studied with this method include morphologically complex words (Silva & Clahsen, 2008), ambiguity and context (Eddington & Tokowicz, 2012; Schwartz & Arêas da Luz Fontes, 2008), and multilingual

naming (Runnqvist, Strijkers, & Costa, 2019). Despite this existing precedent in the literature, semantic priming and lexical decision tasks have yet to be implemented in the syntagmatic and paradigmatic discourse.

Additionally, there are gaps discussed in the literature pertaining to the measurement of L2 proficiency. Wen and Van Heuven (2017) point out that proficiency research often uses dichotomous or discrete categories, such as ‘high proficiency’ and ‘low proficiency’. This limits the ability of researchers to measure its effects on mental lexicon development. Yet, proficiency is not the only limitation in the literature addressed. Wen and Van Heuven (2017) also identify other factors such as language dominance, and age of acquisition, as either absent or understudied.

Among these other factors is the L2 acquisition method. While some studies have looked at techniques and learning strategies used to learn a second language, there are few studies that compare these strategies to each other, particularly for the newer methods, such as subtitles or computer learning (Mohsen, 2016; Yanguas, 2012). Furthermore, there are fewer articles that discuss the learning background of L2 speakers in great detail, which leaves out crucial information about how different techniques or acquisition patterns can change outcomes in lexical organization. Sabourin and colleagues (2014) indicate that future research should also consider the method by which bilinguals acquired their L2, such as in-school, home, or online acquisition, indicating that the method of acquisition may have a role in the development of the bilingual mental lexicon.

Current Study

The current study used semantic priming to examine paradigmatic and syntagmatic associations within the L2 mental lexicon, in order to better understand L2 development when compared to L1. Furthermore, this study examined how various participant factors can influence these associations, with a special focus on L2 proficiency as a continuous variable, and the medium of L2 acquisition. For this study, I recruited participants with English as an L1 and English as an L2. L1 speakers were further divided into Monolinguals and Bilinguals, in order to account for any effects on L1 associations caused by additional languages.

Participants from the three language groups were given a language profile questionnaire and then asked to complete a lexical decision task (LDT) where English words and non-words were presented. The task had four different primes: the two associative primes, clang (or phonetic) primes, and unrelated primes. The participants were first briefly shown a prime and then a target word and were required to respond as to whether the target was a legitimate word or a non-word. Response times and accuracy were recorded for each of the participants' answers. Participants also completed a word association task (WAT) where they were presented a word from a priming list not previously given to the participant, and then given time to type in the first word that came to mind. This task was included both to potentially compare responses to previous work and to check for the validity of the primes selected for the LDT.

It is important to address the circumstances surrounding the administration of the study. Study creation and data collection occurred during the COVID-19 pandemic and its associated restrictions. This meant that in-person administration of the study was not possible, although it did present a research opportunity. By running the LDT online, a broader participant pool could

be accessed, potentially leading to greater applicability of the findings. Therefore, it was important to ask if online administration can effectively record the effects of semantic priming.

Overall in measuring this semantic priming, the hypotheses address whether the organizational patterns of L2 in bilinguals would match those of L1 in monolinguals and bilinguals. While some components of this research were exploratory, there were a few predictions that could be made about the potential results:

1. Online Testing of a Lexical Decision Task is adequate for recording semantic priming.
2. Syntagmatic priming will lead to relatively greater facilitation in a lexical decision task for L2 speakers when compared to L1 monolingual and bilingual speakers. Conversely, L1 speakers will show relatively greater facilitation by paradigmatic priming in a lexical decision task when compared to L2 speakers.
3. Semantic priming will reveal a trend of shifting from syntagmatic to paradigmatic associations driven by L2 proficiency. This means less proficient L2 speakers will favour syntagmatic facilitation in a lexical decision task, whereas more proficient L2 speakers will have greater paradigmatic facilitation in the same task.
4. Immersive methods of L2 acquisition will result in stronger L2 lexical associations in the mental lexicon than non-immersive ones. This means that those who learned with immersive methods will display greater facilitation from paradigmatic priming, akin to L1 speakers. Here, ‘immersion’ refers to the prevalence of L2 in the learning environment, with immersive methods having high prevalence of L2 (e.g., home, formal immersion schooling, exchange

programs), and non-immersive methods having low prevalence of L2 (e.g., online language apps, school or work courses).

5. Monolingual speakers and L1 bilingual speakers will show no significant differences in facilitation effects.

Some of these predictions are also visually depicted in the following graphs.

Figure 1

Hypothesis 1. Facilitation Effects of Associative Primes across Native Language. Note that the degree of facilitation is only representative of direction and not the relative size of the effect.

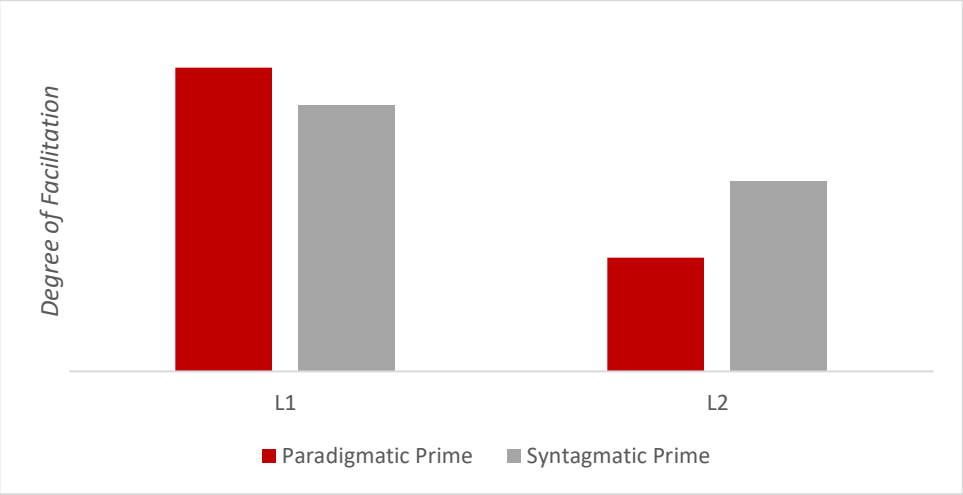


Figure 2

Hypothesis 2. Degree of Facilitation of Associative Primes in L2 Speakers across Proficiency levels. Note that the degree of facilitation is only representative of direction and not the relative size of the effect.

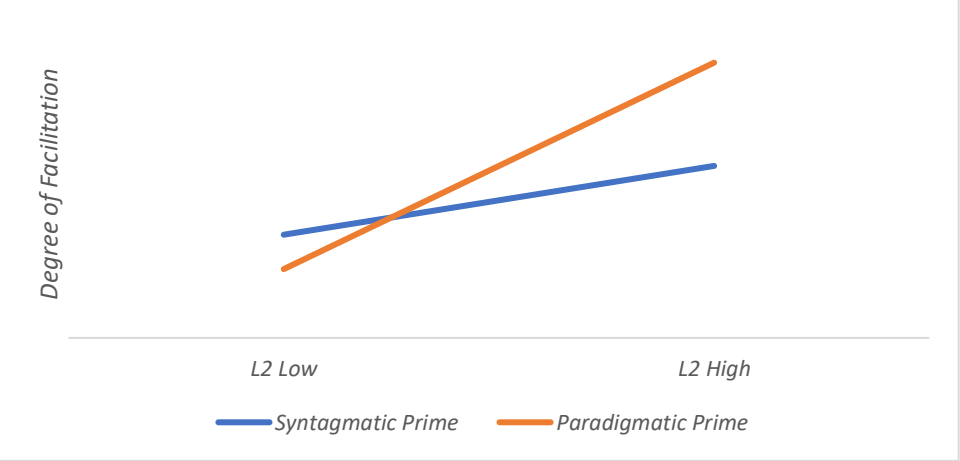
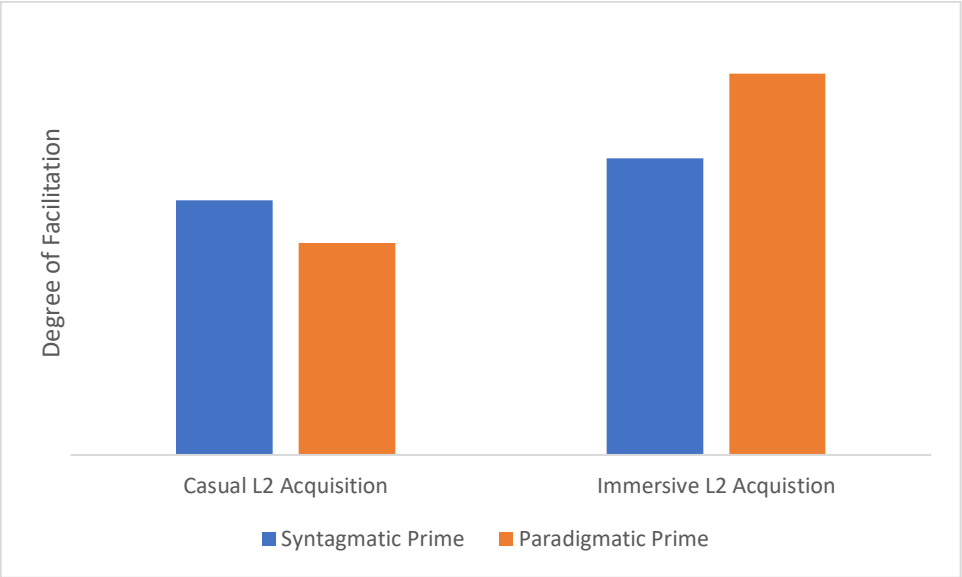


Figure 3

Hypothesis 3. Degree of Facilitation of Associative Primes in L2 Speakers across Learning Methods. Note that the degree of facilitation is only representative of direction and not the relative size of the effect.



Rationale

In order to get a representative sample of participants, the initial goal of the main study was to recruit approximately 80 participants, based on the numbers of participants found in previous studies such as Zareva (2007), Clark (2012), and Lee and colleagues (2018). However, the previous research had either focused solely on bilinguals (e.g., Clark, 2012; Lee et al., 2018), or on the broad differences between monolinguals and bilinguals (Zareva, 2007), meaning they could recruit fewer participants in order to measure effects in their analyses. In the present study, in addition to comparisons between monolinguals and bilinguals, bilinguals were also divided into English L1 Bilinguals, and English L2 bilinguals. Therefore, a larger overall sample size would be required to produce measurable results, with each group consisting of approximately 50 – 100 individuals. The first group included bilinguals whose L1 is English, the second group were bilinguals who speak English as an L2, while the final comparison group had monolingual

L1 English speakers. This inclusion of monolingual anglophones stems from the online format of the study, and the strong likelihood that monolinguals will choose to participate for monetary incentive, regardless of target demographic. Rather than ignore this monolingual demographic, their participation acted as an additional comparison for both bilingual groups.

Some of the previous studies, such as Lee and colleagues (2018) and Clark (2012), focused their research on bilinguals of two specific languages. However, there is less literature that examines how lexical organization interacts between multiple languages. Due to the lack of research about the syntagmatic-paradigmatic phenomenon in the current literature, this study explores speakers in multiple languages, with the goal of revealing possible differences between languages with regards to this phenomenon.

Given that this thesis was conceived during the COVID-19 pandemic, it became clear that in-person testing would not be feasible in order to complete the current study. Therefore, online models of testing were selected as the preferred method, even after pandemic restrictions were lifted and in-person testing began anew. This is because several benefits were identified to be beneficial to the study's needs. First, it allowed for increased recruitment of participants, due to the lack of geographical limitations, particularly crucial in acquiring L2 bilingual speakers. Second, it allowed this recruitment to occur over a relatively short period of time, as there were no laboratory space constraints.

The current study is designed using *Testable* and is hosted on the *Testable* website. However, as mentioned previously, the participants were recruited using *Amazon Mechanical Turk*. During experiment creation, a question arose pertaining to masking the primes. Some studies have argued that when examining "qualitative primes" such as syntagmatic and paradigmatic associations, that the length of the ISI does not significantly affect outcomes

(Zareva, 2007), and there was consideration to have no mask. However as mentioned previously, others argue that in lower proficiency or "unbalanced" bilinguals, short ISI length will result in no effect of priming (Lee et al., 2018). Due to precedents established in the literature, as well as anecdotal testing with laboratory associates, it was thus determined to retain masking, and that 100ms was the optimal length for perception by participants regardless of proficiency.

Furthermore, while the initial aim of this research was to move away from word association paradigms, it became evident that a word association task could be used to verify whether participants would report similar associative responses to the prime-target pairs in the LDT. This could then, in turn, also be compared to the existing literature, to see if the study yields comparable results. Ultimately, the Word Association Task devised was shorter than the LDT in order to reduce fatigue and increase retention amongst participants.

Methods

Participants

Pilot Study. The project was divided into two phases. The first phase was referred to as the pilot study and was conducted in April of 2022. This phase was administered to online participants in order to test the study's stimulus list, to see whether the study was recording responses, and to verify its capacity to record the dependent variables accurately. The participants in the pilot study were all recruited using *Amazon Mechanical Turk*, specifically restricted to North American servers (i.e., Canada and the U.S.). The pilot group consisted of 26 people, 23 of which were bilingual or multilingual. Of these, 14 participants were male and 12 were female, spanning in age from 21 to 54. Fourteen languages were represented, with the most common apart from English being Spanish (5), Korean (3), and Vietnamese (3).

Main Study. Upon completion of preliminary testing, the scope of research was expanded by dropping server location restrictions, resulting in over 259 participants submitting results. These participants also engaged with the study via Amazon *Mechanical Turk* or using *Testable Minds*. A total of 47 participants received an accuracy score lower 70% and were excluded from the results. The removal of this number also included participants who improperly completed the WAT. There was also one participant removed for duplicate testing. Of the remaining 212 participants, 94 were female and 118 were male, with an age range of 19 to 72. Twenty-seven languages were represented in the bilingual participant pools, with the largest apart from English being Spanish (43), Portuguese (14), and Tamil (13). A significant number of participants (94) were classified as "L2 not specified", as they did not disclose that linguistic information.

Stimuli

The Lexical Decision Task (LDT) included six stimulus conditions: two nonword+prime trial types and four word+prime trial types (see Table 1 for examples). The four word-trial types consisted of three with associative primes and one with a prime unrelated to the target word (e.g., apple→ *sky*). The latter category was used to establish a response baseline. In addition to the already mentioned syntagmatic and paradigmatic primes, the third was 'Clang' primes. Clang associations are words that are phonologically and/or orthographically linked to a target word, rather than by semantic or syntactic associations, e.g., rhymes (i.e., *dog* → *log*, *axe*→ *lacks)* or homonyms (i.e., there → their) (Clark, 2012). This priming setup was chosen due to previous precedent in the literature (Fitzpatrick & Thwaites, 2020). In total, there were 160 word stimuli and 160 non-word stimuli compiled for this project. The word stimuli were all nouns, primarily concrete nouns, compiled partially from the stimuli from previous experiments (Clark 2012; Sabourin et al., 2014) and from lexical databases, such as the *Leipzig Corpora Collection*. Due to the use of North American online databases for participants, when target words had British and American variants in spelling, the American spelling was used.

Once compiled, each target word was assigned a prime from each of the four priming categories. Here participants would only see a target word once, accompanied by one type of prime. For prime selection, potential paradigmatic primes were restricted to concrete nouns in the same category as the target words, with exceptions only when a concrete noun association was non-existent. Syntagmatic primes selected for the study included concrete nouns, simple verbs, and simple adjectives related to the target prime. "Clang" primes were restricted to words containing a phonological or a phonological-and-orthographical relationship to the target word and were compiled using an online rhyming dictionary. If the initial rhyme was too uncommon

(e.g., technical jargon, archaic spellings, etc.) or unavailable, an approximate rhyme was used instead. Unrelated primes consisted of concrete nouns with no semantic or associative connection to the target words, all drawn from an online random word generator to reduce bias. Once the initial list of primes was completed, it was reviewed for potential confounds. During this process, if a prime exhibited associative ambiguity (i.e., could function as more than one association type), the word was either altered from root form (e.g., *ring* → *ringing*) or substituted with another prime. If a prime could be used for multiple target words, it was either restricted to one target, or discarded entirely. Table 1 illustrates a sample of word stimuli used in the study (for a full list of targets and primes, see Table 6 in Appendix A).

Table 1

Sample of Target words and their associated primes

Target Word	**Syntagmatic Prime**	**Paradigmatic Prime**	**Clang Prime**	**Unrelated Prime**
ANCHOR	marine	boat	banker	tame
APPLE	cider	pear	grapple	tomb
AXE	grind	sword	lacks	libel
BEACH	sandy	park	leech	lye
BELL	ringing	chime	spell	scale
BELT	leather	shirt	melt	collect
BOOK	read	magazine	hook	race
BREAD	butter	toast	thread	basil
BRIDGE	suspension	overpass	fridge	common
CAKE	birthday	pie	flake	disco

In the case of Non-Word Targets, an additional list of 160 targets was created using *Wuggy*, a computer application designed to create randomized lists of non-words from a word input. First, word inputs were selected from the online random word generator, and then processed via *Wuggy*. The software would then generate a list of potential non-words derived from the word input. Two or three non-words were selected from the generated list, and the

process would repeat until 160 non-words were chosen. Upon further review, the non-word targets were examined for their conformity to English grammar and word length compared to word targets. If a non-word phonetically sounded like an English word, did not conform to English spelling, was over 7 letters long, or had spelling with ambiguous pronunciation, it was either modified to conform to these parameters or eliminated. Since non-words cannot form grammatical associations, only clang and unrelated primes were compiled using the same methods as previously used for the Word List. Table 2 illustrates a sample of non-word stimuli used for the LDT; for a full list of targets and primes, see Table 7 in Appendix A.

Table 2

Sample of Target non-words and their associated primes

Target Non-Word	Clang Prime	Unrelated Prime
CHEMMY	tummy	moral
CLARD	hard	daughter
CLEATH	sheath	unique
COTEE	goatee	request
COUDAY	holiday	jail
CRIX	tricks	mold
CROOM	loom	ditch
CROTH	sloth	colon
DIMIN	ermine	fix
DIPPANT	flippant	waist

For counterbalancing purposes, there were four stimulus sequences created. The stimuli were divided into four 320-prime test lists, consisting of all the targets each with a prime from one priming category. Each list contained an equal number of primes from each category. For example, the target word “ANCHOR” would be paired with the syntagmatic prime “marine” on List 1, while the following target word “APPLE” would be paired with the paradigmatic prime “pear” on the same List. As demonstrated in the previous example, targets were presented in

upper-case letters, while primes were presented using lower case letters, to avoid orthographic confounds. This cycle would also occur for the non-word targets, though it would be restricted to clang and unrelated primes only. Table 3 provides an example of this layout below.

Table 3

Sample of Counterbalanced Prime Lists for Participants

Target Word	**Participant List 1**	**Participant List 2**	**Participant List 3**	**Participant List 4**
ANCHOR	marine	boat	banker	tame
ALK	well	chalk	well	chalk
AXE	lacks	libel	grind	sword
ARLIK	bootlick	service	bootlick	service
BELL	ringing	chime	spell	scale
BALOR	real	valor	real	valor
BOOK	hook	race	read	magazine
BELUE	venue	flex	venue	flex
BRIDGE	suspension	overpass	fridge	common
BLUTE	unit	flute	unit	flute

For the word association task, no additional stimuli were created. Rather, the participant received half of one of the priming lists that they had not seen during the LDT. For example, if participant had been assigned List 1 during the LDT, they might be assigned half of List 2 during the word association task.

In addition to the priming lists, two questionnaires were used for the participants to answer prior to the study. The first questionnaire asked participants questions related to their bilingualism, such as their age of L2 acquisition, daily language usage, places and people associated with L2 use, acquisition methods, among other relevant data. The purpose of the questionnaire is to develop a detailed language profile for participants, addressing one of the challenges described by researchers about collecting sufficient information related to speakers' languages to construct continuous variables of language fluency and acquisition (Van Hell &

Tanner, 2012; Wen & Van Heuven, 2017). In addition to linguistic information, participants' demographics were also recorded in order to rule out potential compounds.

The second questionnaire is an English Proficiency test, previously used in the Words in the World Lab. The questionnaire consists of twenty questions relating to specific daily tasks and conversations, asking participants to rate their proficiency in each circumstance separately. The purpose of this study is to provide an in-depth view of participants' proficiency, which has been often limited by vague questions in self report (e.g., *rate your proficiency on a scale of 1-10*). Furthermore, by combining these results with the linguistic profile, it allows us to better examine the differences between different bilinguals in addition to the oft-examined monolingual-bilingual difference. Figures 4 and 5 show a sample of the questions asked in both questionnaires, with the full list of questions found in Appendix B.

Figure 4

Sample of Linguistic Profile Questions

3. How long have you been speaking a second language?
    a. I do not speak a second language
    b. Less than a year
    c. 1-2 years
    d. 3-5 years
    e. 5-10 years
    f. 10+ years
4. How did you first learn a second language? (Select all that apply)
    a. I don't speak a second language
    b. From family
    c. From classes
    d. At work
    e. From printed media
    f. From broadcast media
    g. From social media
    h. Other (please specify)

Figure 5

Sample of English Proficiency Questions

I can explain why a certain location is my favourite vacation spot by providing three reasons.	1 = I cannot do this
I can refuse an invitation to go somewhere (for example, to a restaurant) and provide reasons for my refusal.	2 = Poorly
I can say what I did yesterday.	3 = With Difficulty
I can start a conversation when I meet someone.	4 = Reasonably Well
I can return an item to a store and explain why I am returning that item.	5 = I can do this fluently

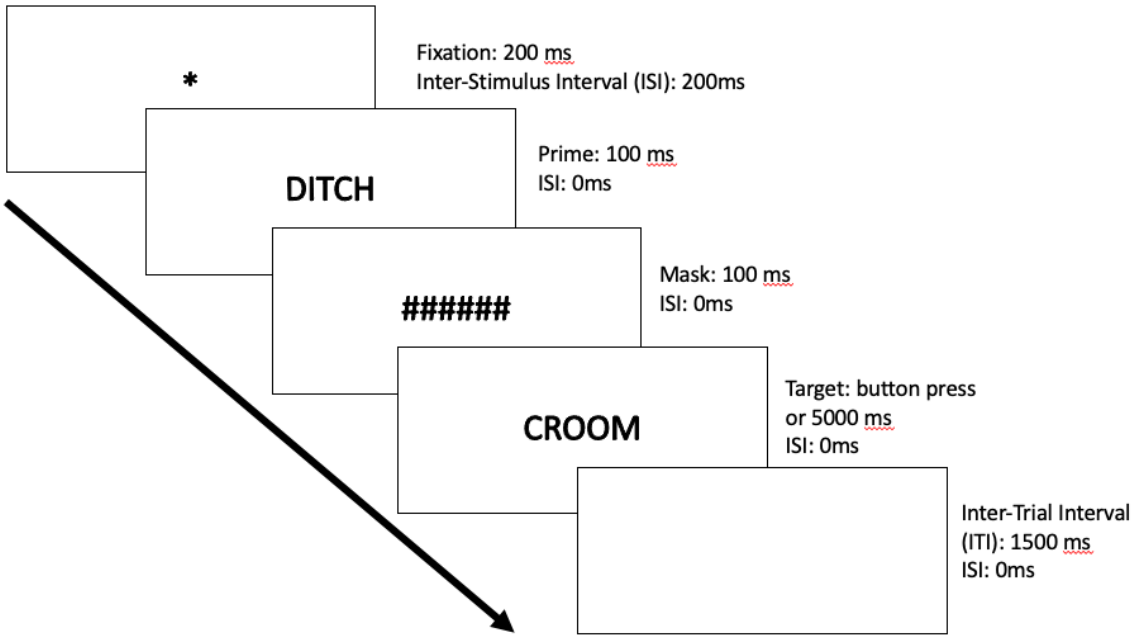
Procedure

The study procedure was divided into three sections: Questionnaires, LDT, and WAT. First, participants filled out the questionnaires prior to the start of the study. Upon completion, they began the LDT, where the participants were shown a fixation stimulus (*), a prime, a mask (######), and a target. The participants were required to respond whether the target was a word or non-word using the “A” and “L” keys, respectively. If they pressed the wrong key, a sign indicating “incorrect response” would appear for one second, to discourage random button pressing. Their answers were recorded as “0” if false, or “1” if correct. There was a 200ms ISI between the fixation and the prime, followed by displaying the prime for 100ms. The mask immediately replaced the prime for 100ms, before showing the target until a response key was pressed or for a maximum of 5000ms. There was no ISI between the prime, mask and target, but

once the trial was complete a blank screen would signal the start of the ITI of 1500 ms, and the cycle began again. Figure 6 illustrates this full sequence of a trial in further detail. This process continued until the participant has completed 320 randomized trials, with half the trials being non-words. Response times and accuracy were recorded, for each of the participants' answers. Following the Lexical Decision Task, the participants were given a 30 second break, and a follow-up question, asking them to rate their attention levels during the task using a Likert scale from 1-10.

Figure 6

Presentation of Stimuli in the Lexical Decision Task (LDT)



Participants were then presented with the word association task, in which they were presented a word for 500ms, and then given time to type in the first word that came to mind. During piloting, this period lasted 15 seconds, and in the main study, it was extended to 60

seconds. If a participant took longer than the allotted time to complete a response, or did not respond, it was recorded as a timeout, and the next word appeared. While theoretically it was possible to let the response window timeout for every trial, only one participant elected to not respond to the WAT. This task consisted of 80 trials, followed by another attention check using the same 1-10 Likert scale, thus finishing the study. Participants took an average of 30 minutes to complete the whole study, with the shortest timeframe being 20 minutes and the longest being 50 minutes. This variation arises from the mandatory nature of responses during the LDT, resulting in some respondents taking breaks throughout the task. Upon completion, participants received compensation, initially set at $2.00 CDN per person, but later increased to $4.00 CDN when examining the time participants spent performing the study tasks.

Analyses

The analyses were conducted using Microsoft Excel and R. Linear Mixed Effect Regression (LMER) analyses were conducted including several covariates recorded by the questionnaire, such as age, gender, education level of the participants, accuracy of their performance, target word length and frequency of the word primes and target words.

The reasons for using R are as follows. Standard statistical analysis comparing groups or tasks in cognitive psychology has involved ANOVA designs. For example, in a working memory task comparing 1-back performance (pressing a key when the digit currently presented is the same as the one immediately previous) to 2-back performance (pressing a key when the digit currently presented is the same as the one 2 trials back), a simple ANOVA design is adequate because the stimuli are all the same (digits 1 to 9) and comprise essentially the entire population of stimuli of interest.

However, in psycholinguistic studies with words as stimuli, the words are sampled from the extremely large potential population in the language and therefore should be treated as a random variable, just as research participants are. Furthermore, the words have characteristics that influence reaction times, most notably frequency of occurrence in the language and word length. In order to enter word-characteristics as a covariate, the stimulus conditions (in our case, the nature of the primes and the targets being used in each trial) must be treated as a random variable. Only LMER permits this. For this reason, LMER has become the default statistical procedure for use in psycholinguistics. Note that with LMER, results are given in terms of the t-value significance of the beta weights, rather than F-scores found with ANOVA designs. Furthermore, in LMER the data from all acceptable trials are included in the analysis, in the present case totalling 211 participants x 160 word trials = 33,920, yielding potentially very large degrees of freedom. In all analyses run in this study, the degrees of freedom exceeded 5000, and thus differences across analyses are inconsequential. For this reason, *t*-values are given without degrees of freedom indicated.

The results analysed were filtered to look at correct responses only, and only for participants scoring more than 70% in overall accuracy. This reduction led to an 81.85% retention of data. Furthermore, the remaining data were trimmed to exclude reaction times (RT) greater than 2000ms, and less than 300ms, which resulted in an additional 2% reduction. Furthermore, two sets of analyses were performed, one where the independent variable (IV) Prime Type was measured against native language (i.e., monolinguals vs. bilinguals) and the other where the IV was measured against Linguistic status (e.g., L1 Bilinguals). The first LMER analysis included several covariates, namely demographics (i.e., age, gender, and education), target word frequency, target length, previous reaction time, and trial number. Participant

accuracy was added on subsequent runs of this analysis. The second LMER analysis maintained these same covariates; however, when isolating the language groups for separate follow-up, only participant accuracy, age, gender, and word frequency were maintained, as the other factors did not have an effect on the dependent variable. Due to this large number of covariates and precedent in LMER analyses, alpha level was adjusted to $p = .01$, rather than .05 used in many psychological analyses.

Results

Pilot Study

The findings yielded no significant issues requiring attention or modification, thus allowing for the research to proceed to the main study. In the lexical decision task, it was determined that participants responded significantly faster to words than to non-words (764 ms vs. 842 ms; $t = 2.10$, $p = .04$), in line with previous research using the paradigm. Furthermore, only four stimuli were identified as near or below chance levels (2 non-words, 2 words). These discrepancies may have arisen from homophones to the targets (e.g., PAIL and 'pale', OAR and 'or' / 'ore'). The low number of stimulus issues accounted for less than 1% of the total stimuli; given the limited amount of time available for pilot testing, the decision was made to retain these primes in the main study. For more information about stimulus accuracy in the pilot study see Tables 8 and 9 in Appendix C.

Main Study

Three separate analyses were conducted to test for facilitation effects on reaction times in the Main Study. These models primarily compared groups by specific language status: by native language (L1 vs. L2), comparing native English-speaking monolinguals and bilinguals, and L1 versus L2 bilinguals. These 2x2 analyses were run in order to specifically address hypotheses 1, 2, and 4. General description of response and facilitations times is available in Tables 4 and 5 below, with a visual representation of the facilitation effects shown in Figure 7.

Table 4

Mean Response Times for Lexical Decision Task by Prime Type and Participant Language Status. Analyses were run with all participant groups included. Means have been adjusted for covariates. SE = Standard Error; asymp.UCL & LCL = Asymptotic Upper & Lower Confidence Limit at .01

Participant Language Status	Prime Type	Response (ms)	SE (log)	asymp.LCL	asymp.UCL
L1 Bilingual	UNRELATED	862.1	0.02	825.9	899.9
L1 Monolingual	UNRELATED	844.6	0.03	796.4	895.7
L2 Bilingual	UNRELATED	756.8	0.03	714.0	802.2
L1 Bilingual	CLANG	859.6	0.02	823.5	897.3
L1 Monolingual	CLANG	847.4	0.03	799.1	898.6
L2 Bilingual	CLANG	743.1	0.03	701.1	787.6
L1 Bilingual	PARADIGMATIC	850.0	0.02	814.3	887.3
L1 Monolingual	PARADIGMATIC	834.4	0.03	786.4	884.8
L2 Bilingual	PARADIGMATIC	736.0	0.03	694.4	780.1
L1 Bilingual	SYNTAGMATIC	845.8	0.02	810.3	887.9
L1 Monolingual	SYNTAGMATIC	829.2	0.03	781.8	879.3
L2 Bilingual	SYNTAGMATIC	739.8	0.03	697.9	784.1

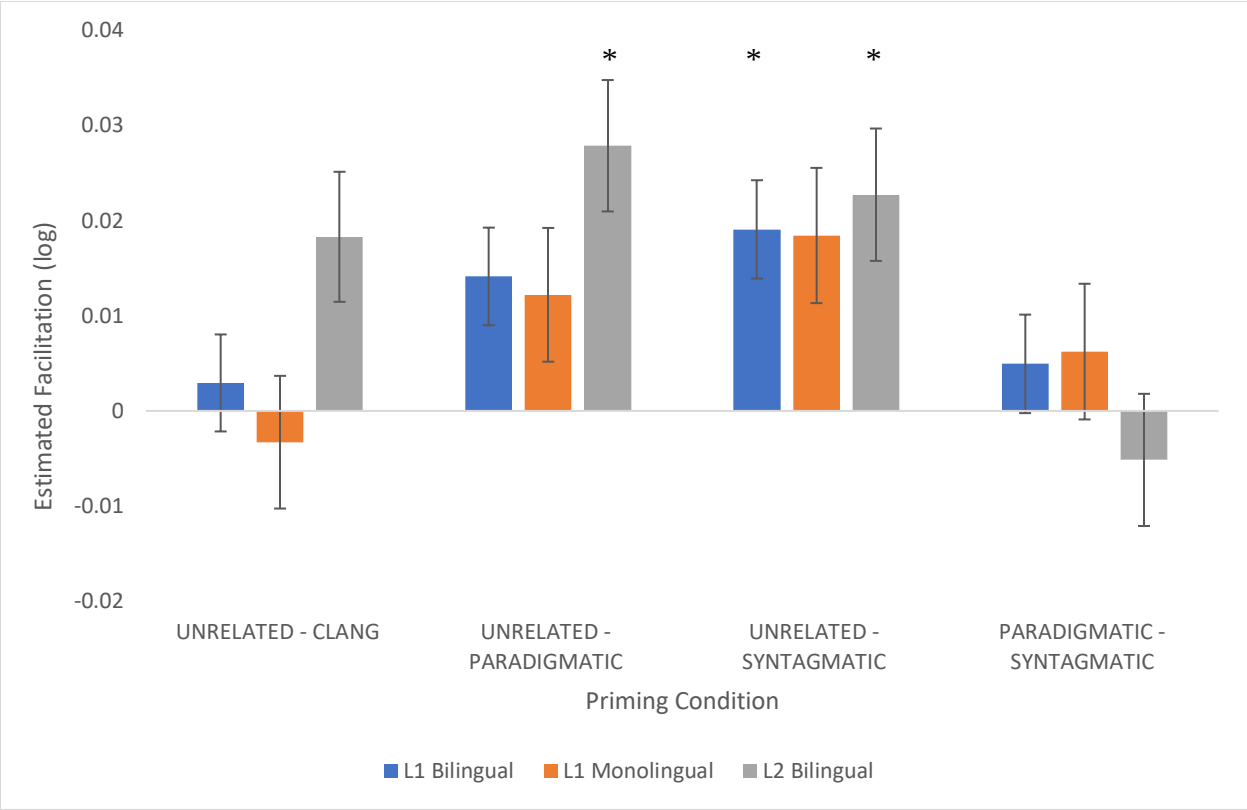
Table 5

Facilitation Effects for Lexical Decision Task by Participant Language Status. Analyses were run with all participant groups included. Statistically reliable effects are indicated by an asterisk.

Participant Language Status	Prime Comparison	Facilitation Effect (ms)
L1 Bilingual	UNRELATED – CLANG	2.5
L1 Bilingual	UNRELATED – PARADIGMATIC	12.1
L1 Bilingual	UNRELATED – SYNTAGMATIC	16.3*
L1 Monolingual	UNRELATED – CLANG	2.8
L1 Monolingual	UNRELATED – PARADIGMATIC	10.2
L1 Monolingual	UNRELATED – SYNTAGMATIC	15.4
L2 Bilingual	UNRELATED – CLANG	13.7
L2 Bilingual	UNRELATED – PARADIGMATIC	20.8*
L2 Bilingual	UNRELATED – SYNTAGMATIC	17*
L1 Bilingual	PARADIGMATIC – SYNTAGMATIC	4.2
L1 Monolingual	PARADIGMATIC – SYNTAGMATIC	5.2
L2 Bilingual	PARADIGMATIC – SYNTAGMATIC	3.8

Figure 7

Degree of Facilitation between Participant Language Categories across Priming Condition. Y axis measures facilitation logarithmically.



Native Language

When analyzing the findings using the binary variable of native language, there is a reported main effect of native language, with L2 speakers, somewhat paradoxically, responding in general more quickly than L1 speakers ($t = 5.29$, $p < .001$). Furthermore, there was a main effect of prime type indicating overall facilitation for Paradigmatic primes ($t = 2.70$, $p = .007$) and Syntagmatic primes ($t = 3.82$, $p = .0001$) compared with unrelated primes, with no significant difference between Paradigmatic and Syntagmatic facilitation ($t = 1.14$, $p = 0.25$).

Following up this analysis, LMER analyses were run for L1 and L2 speakers separately (see Tables 14 and 17). It was found that syntagmatic and paradigmatic primes showed shorter

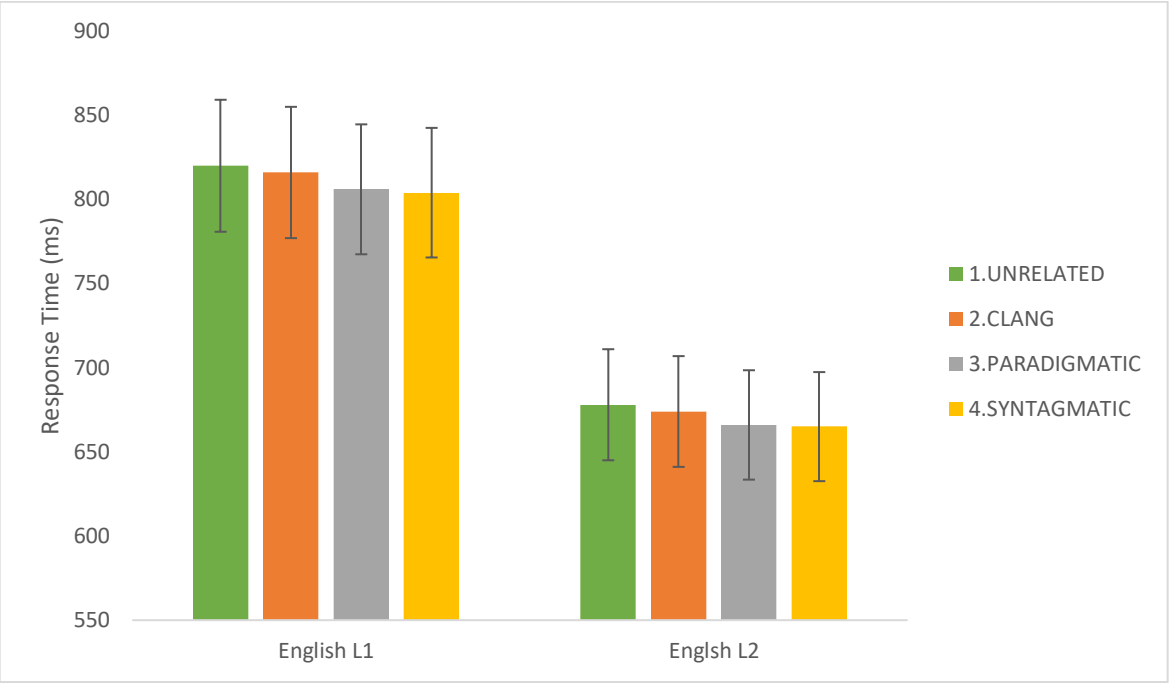
response times compared to unrelated primes in both L1 speakers ($t = 3.21$, $p = 0.001$; $t = 4.39$, $p < 0.001$) and L2 speakers ($t = 4.04$, $p < 0.001$; $t = 3.44$, $p < 0.001$). Examined together however, two-way interactions comparing Groups did not reach significance for syntagmatic priming ($t = 1.48$, $p = 0.14$), but did reach significance for Paradigmatic priming ($t = 2.51$, $p = 0.01$).

For native speakers of English, clang primes did not show response times that differed from those of unrelated primes ($t = 0.24$, $p = 0.81$). In the L2 Bilingual analysis, however, RTs to words preceded by Clang primes were marginally faster than those that were preceded by Unrelated primes ($t = 2.43$, $p = 0.02$). There was also a marginal interaction for Group x Priming conditions for Clang priming (t=2.31, p=0.02), although in this instance the actual difference in facilitation was so minimal as to not warrant follow-up.

In each of these instances, L1 Speakers showing greater facilitation than L2 Speakers. Figure 8 illustrates the main effects of language status, prime type, and their interaction, with reaction times recorded in milliseconds. For full information about these analyses, refer to Tables 12, 14, and 17 in Appendix C.

Figure 8

Effects of Prime Type on Reaction Time measured by Native Language. Error bars represent Standard Error (SE)



Additionally, analyses were run with participant accuracy added as a predictor of reaction time. These analyses revealed that the only significant interaction was between syntagmatic priming and linguistic status (t = 2.67, p = .02) where participants with higher accuracy scores (i.e., showing more fluency), showed greater syntagmatic priming than those with low accuracy scores. Finally, there was the usual main effect of Frequency (t = 11.48, p < .001) with more frequent words being reacted to faster than less common ones.

Participant Language Status- Monolinguals & L1 Bilinguals

Comparing Estimated Marginal means between the participant groups revealed that there was no significant difference between English L1 Bilinguals and Monolinguals with respect to overall reaction times (z = 1.20, p = 0.45) as shown in Table 10. Further LMER results indicate

no significant difference in overall facilitation between these two groups ($t = 0.18$, $p = 0.85$). By itself, the L1 Monolingual group only reports an effect of syntagmatic priming ($t = 2.72$, $p < 0.01$), and only a marginal effect of paradigmatic priming ($t = 1.72$, $p = 0.07$). This slightly contrasts to these conditions in the L1 Bilingual group, which reports main effects for both syntagmatic ($t = 3.48$, $p < 0.001$) and paradigmatic priming ($t = 2.63$, $p < 0.01$). Neither Monolinguals ($t = 0.41$, $p = 0.68$) or Bilinguals ($t = 0.53$, $p = 0.60$) report main effects of clang priming. However, the differences did not result in an interaction between Group and Priming Condition.

Main effects of frequency were consistent with the previous analysis ($t = 10.76$, $p < .001$). For further details, see Tables 13, 15, and 16 in Appendix C.

Participant Language Status – L1 & L2 Bilinguals

Main effects of frequency were consistent with both previous analyses ($t = 11.20$, $p < 0.001$), with less frequent words resulting in longer reaction times. Reaction times were also affected by participant language status, as described above for the overall analysis: L2 Bilinguals responded more quickly than L1 Bilinguals in general ($t = 5.66$, $p < .001$). However, none of the 2-way interactions of Group x Priming conditions reached significance indicating no group differences in facilitation effects.

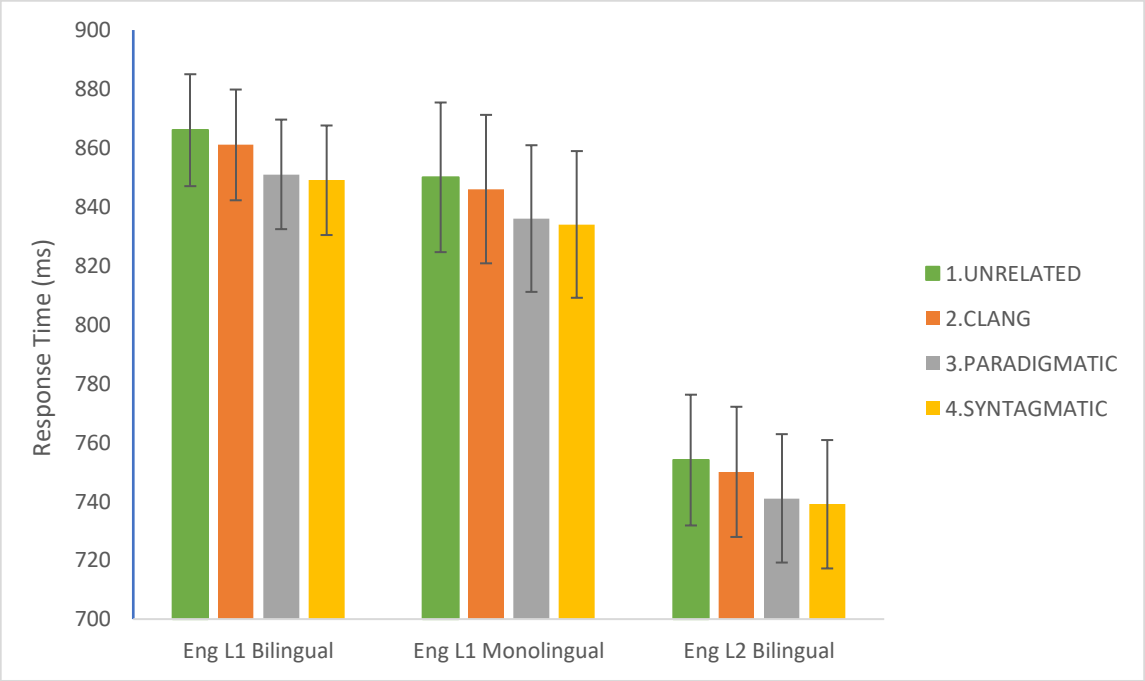
The L1 Bilinguals in isolation reveal similar findings to the Native Language analyses, with significant facilitation effects for paradigmatic priming ($t = 2.63$, $p < 0.01$) and syntagmatic priming ($t = 3.48$, $p < 0.001$). Here there are no facilitation effects between clang and unrelated

primes (t = 0.53, p = 0.60) or difference in the size between syntagmatic and paradigmatic priming (t = 0.87, p = 0.38).

Looking at the L2 Bilinguals separately, the facilitation effects of prime type extend to Syntagmatic primes (t = 3.44, p < .001) and Paradigmatic primes (t = 4.04, p < .001), while also indicating marginal effects from Clang primes (t = 2.43, p = .02). Furthermore, the three priming conditions, although showing facilitation in the same direction, do not differ from each other in two-way comparisons: Syntagmatic versus Paradigmatic priming (t = 0.57, p = 0.57), Syntagmatic versus Clang priming (t = 1.06, p = 0.29), and Paradigmatic versus Clang priming (t = 1.66, p = 0.10). For further information, see Tables 14 and 15 in Appendix C.

Figure 9

Effects of Prime Type on Reaction Time measured by Participant Language Status. Error bars represent Standard Error (SE).



Demographic Findings

Apart from the main effects relating to type of prime, there are demographic main effects as well. The findings indicate a main effect of Gender in the whole dataset (t= 2.14, p=.03), with participants who identified as women responding more slowly than participants who identified as men. However, when measuring the linguistic groups separately, this finding disappears in L2 Bilinguals (t = 0.87, p = 0.39) and L1 Monolinguals (t = 0.53, p = 0.60), only appearing marginally in L1 Bilinguals (t = 1.74, p = 0.08). Furthermore, there was a main effect of age in L2 Bilinguals where older L2 speakers responded more slowly than younger L2 speakers. (t = 3.04, p = 0.01). By contrast, there was no main effect of age in the remaining language groups (see Tables 14 – 16 in Appendix C). Interestingly, the mean age by group is oldest for the L1 Monolinguals (40.45 years), with the L2 Bilinguals being the youngest group on average (34.03 years). Despite these effects, the findings do not affect facilitation times whatsoever.

As previously mentioned in the methods section, questionnaire data were collected concerning the methods of language acquisition in order to construct linguistic profiles for each respondent. However, due to the nature of the questionnaire design, the data collected were not easily translatable into quantifiable data. For example, participants often selected every possible response on questions pertaining to acquisition methods, daily use, and frequency. However, because each response was only recorded via checkbox, it meant that all responses were weighted equally, and led to high rates of confounds in the data.

Word Association Task

In addition to the LDT, the participants also did a Word Association Task during the study, with the goal of supporting the LDT results. However, there were unexpected consequences due to the setup of this task. Initial wording of the instructions led many participants to simply type the presented word in the response box. Despite clarification to the instructions, this issue persisted throughout the experiment, leading to a small data set via exclusions. The constraints of the data, and potential participant fatigue led to the decision to not analyse the data for the purpose of this thesis.

Discussion

The purposes of this study were to examine syntagmatic and paradigmatic associates in the L2 mental lexicon and to compare the findings to the L1 mental lexicon. When comparing the findings to the original predictions, several conclusions emerge. First, the findings support Prediction 1, with the online LDT adequately recording differences in semantic priming. This is evident in the pilot data that shows faster reaction times to words over non-words, and in the main study, where overall priming of syntagmatic and paradigmatic primes result in faster LDT responses than unrelated ones. The findings also reveal significant facilitation of paradigmatic and syntagmatic priming in L2 speakers. The findings further indicate that L2 bilinguals did not favour syntagmatic primes over paradigmatic ones, as predicted in Hypothesis 2. The study further revealed a robust facilitation effect only from syntagmatic primes in L1 speakers, rather than the expected greater facilitation for paradigmatic primes.

While Prediction 3 was not tested as rigorously as the previous ones, the findings do not support it either, but rather the reverse, with low proficiency L2 speakers showing more paradigmatic priming and high proficiency L2 speakers showing more syntagmatic priming. Prediction 4 could not be verified nor disproved at this time due factors of design and communication, which will be discussed further in the limitations section. Despite the lack of support for predictions 2 to 4, the findings were consistent with prediction 5, with no significant difference between L1 monolinguals and L1 bilinguals in terms of reaction times and facilitation. Above all, the findings reveal that the two measured categories of associative primes do not significantly differ from one another. The findings thus support other work, such as Zareva (2007), that indicate no significant difference between association types. However, there are ambiguities that arise from these findings that do not allow for simple conclusions to be made,

partially due to a need for more data.

Despite the findings not supporting the hypotheses, the exact way they differ may be a subject of further inquiry. Previous research has indicated that L2 speakers would react slower when compared to native speakers (Zareva, 2007). However, our results indicate that L2 speakers reacted faster in the LDT compared to L1 speakers across all priming categories. This has also been found to be the case sometimes by other researchers (e.g., Sabourin et al., 2014). One possible reason for this occurrence may be that the larger size of the native speakers' English mental lexicon compared to the L2 speakers. For example, Segalowitz and Zheng (2009) demonstrated that when the list of words in a block were from a single category (e.g., animals) for a LDT, participants reacted more quickly than when stimuli were not semantically grouped and could be from one of 5 categories. We would expect that in most L2 speakers, there is a narrower lexical window when compared to L1 speakers, perhaps increasing the speed at which they can react in the LDT.

Mental Lexicon Formation

The findings support the idea that the L2 mental lexicon does not fully resemble that of participants who have English as their L1. The tendency of L2 speakers to be more facilitated by paradigmatic primes than native English speakers suggests that there are differences in linguistic associations, although these are not radical departures from L1 speakers over the age of 8, consistent with work of previous researchers of the Syntagmatic-Paradigmatic shift (Ervin, 1961).

Alternatively, these results could also emerge from the word choices employed in the study. As mentioned before, examples of stimuli for associative priming in the literature have been simple concrete nouns (Sabourin, 2014), which was emulated and repeated here. However,

these stimuli can also be some of the earliest words learned by English L2 speakers, an example being used vocabulary lists in learning apps (Teske, 2017). By using words that are often learned paradigmatically or phonetically, the likelihood of paradigmatic association in this group increases. Likewise, the marginal association of Clang primes in L2 speakers could also arise as a result of this preference (although the extremely effect small size suggested by our data does cast doubt on its replicability). The findings point to the possibility that if early learning of L2 uses different types of association in the lexical storage system (e.g., more syntagmatically), then it could lead to greater facilitation of those primes early in the L2 learning process.

Another possibility arises from Zareva (2007), where it is mentioned that L2 speakers with larger vocabulary sizes led to greater facilitation and heterogeneity in associative connections. According to these results, this means that the more word knowledge obtained, the more likely an L2 speaker would store associations just like an L1 speaker. However, these findings were obtained using word association paradigms, which are measures of cognitive processing rather than automaticity, which means the types and degree of facilitation may differ. While our findings initially reflected this process occurring in L2 speakers, they did not persist in the data after recruiting a greater number of this language grouping. However, the findings do not rule out the possibility that initial learning of the participant sets the stage for the development of their L2 mental lexicon, with subtle shifts in organization as more vocabulary is acquired.

The Small Priming Effect Problem

When looking at the effects of priming overall, the results indicate smaller priming effects than expected regardless of prime type. Our facilitation effects only averaged 15-20 ms, which remained constant across language status, though one priming category had a marginal

effect as small as 14 ms. The facilitation effects were negligible in participants with overall accuracy below 90 percent. The general semantic priming literature records facilitation greater than 35 ms (Burgess & Simpson, 1988) and up to 100 ms (Favreau & Segalowitz, 1983). However, this facilitation window is consistent with the findings of other work as well. For example, Sabourin and colleagues (2014) measured a facilitation time between 20-25 ms between unrelated and associative primes, while Libben and de Almeida (2000) reported semantic priming facilitation effects of 15 and 20 ms. One possible reason for such a small facilitation effect could be the length of the priming window. Lee, Kang, and Choi (2018) found interactions between L2 language fluency and interstimulus interval (ISI) length, where shorter windows resulted in smaller priming effects in less proficient bilinguals. In the current study, this was evident in participants with low overall accuracy, with the only departures occurring the highly accurate L1 participants in the Paradigmatic priming condition.

One of the ongoing discussions throughout study creation was finding the best prime-target delay window for participants, due to the range of fluency that was expected from the sample. After anecdotal and pilot testing, an ISI of 100ms was determined to fit the purposes of this study best because we found L1 speakers were beginning to use other processes and lose focus with longer presentation windows, such as 150ms. However, it is possible that the priming effects could be more pronounced by extending the ISI window to 150 or even 200 ms. While this modification would help facilitate priming in less proficient bilinguals, it would begin to add confounds for measuring L1 bilinguals, simultaneous bilinguals, or monolinguals as the long ISI window could allow for other mental processes to activate as well. This introduction of confounds would have to be mitigated by potentially using different ISI windows according to L1 or L2 proficiency. For example, Favreau and Segalowitz (1983) found that different ISI

windows in LDT would result in the activation of different processes, such as inhibition versus automatic processing, and that language fluency determined which ones would activate in L1 and L2. Using a longer window for unbalanced L2 speakers might help clarify whether it is only an issue of processing time, or if other cognitive processes might explain the facilitation anomalies. This finding does highlight the challenge in creating continuous fluency variables in mental lexicon research, as discussed by Wen and Van Heuven (2017).

Motivation & Gamification

One unexpected factor that emerged during data collection was the issue of compensation. There was a noted increase in accuracy scores and retention of participants after the compensation for the study was increased. There were participants who indicated that the initial compensation was insufficient for the time spent on task, an issue which did not appear following the increase. This phenomenon was particularly evident in the L1 Bilingual group; however, that may also be a result of the greater number of participants in that category overall. This change in performance post-monetary increase may indicate that motivation can potentially act as a confound or covariate for language proficiency, which would need to be accounted for in future research using LDT tasks. Another way to look at this change is to look at the participants themselves. Having conducted this study on *Amazon Mechanical Turk* and *Testable Minds*, screening participants using the site filters was created so that high-performing participants were recruited. However, by using these sites, the pool of potential participants is narrowed to those who are computer literate and those motivated by financial compensation. This narrowing is further compounded by our decision to not limit languages measured nor countries included, which can mean the value of financial compensation can range from negligible to extremely significant. Overall, the implicit exclusion criteria of high-performance and pay potentially

remove other bilinguals and monolinguals, who are motivated by internal reasons (e.g., fun, academic excellence, competitiveness) and likely not captured to the same degree by the current recruitment format. Future research may want to include controls for financial motivation by using both online and in-person pools and diversifying recruitment places (e.g., schools, libraries, public spaces, online).

Limitations

As mentioned previously, one of the possible limitations of the study was its online nature. While online research allowed for the gathering of large numbers of participants, there was evidence that several participants did not understand the task, despite screening and selection criteria in the online database. These discrepancies in the paradigm may help account for some of the weak findings. Apart from the psycholinguistic hypotheses proposed in the study, by proceeding with data collection via online means, there were assumptions made that the data recording and presentation will be consistent across all participants. However, not all participants used the same operating system, nor the same screen dimensions. These differences can be magnified if internet connection is poor, hardware processing speed is slow, or glitches appear for participants performing the task. Having some in-person participants would have helped to establish an expected baseline for full performing individuals in the online data. However, due to the ongoing effects of the COVID-19 pandemic, this option was not available at the time of testing.

Another limit of the online method of data collection was the inability to do potential follow-up work. The decision to include the word association task as part of the same study as the LDT was influenced by this limitation and may have negatively affected the results on the former as a result of participant fatigue. Any future effort to re-run the experiment with

participants performing all tasks must consider having some in-person work in order to allow for the study to separate into multiple sessions. Alternatively, a narrowing of the participant pool, such as using solely university students, could allow for this work to remain online.

Furthermore, the accuracy of the self-report of participants' language ability had to be addressed several times. Some participants indicated only speaking languages other than English (e.g., Tamil), despite self-reported and tested English proficiency. In other circumstances, despite providing participants ample opportunity to include all languages spoken, there was the large number of L1 bilingual participants who did not disclose their L2, or they initially claimed monolingual status, leaving their language status ambiguous. This broad range of responses reflects the diversity of views that participants have about their own bilingualism, sometimes requiring multiple examinations of the data in order to create a complete picture about their linguistic profile.

The limitations of the questionnaire were also apparent when looking at the questions about acquisition methods for L2. Some participants would select all possible options of acquisition (technically possible), while others misunderstood the instruction and typed unrelated reasons in the clarification boxes. This high variability made examinations of Hypothesis 3 unfeasible given the data, and would require an overhaul of the question style, or geographic restriction to properly capture the information necessary to measure Hypothesis 3.

Unexpected Results and Issues to Explore

While the research was unable to address all hypotheses, the data that emerged did provide a few unexpected patterns. First, there were unexpected findings specifically related to age for L2 speakers, where increased age lessened the facilitation effect in L2 bilinguals. Since this did not occur in the other language groups, it can point to changes in the L2 mental lexicon

across the lifespan. Despite these leads, it is unclear if the changes in the research arise from later age of acquisition, or age-related changes. Future research would need to factor both age of participants and age of acquisition into profiles in order to measure these changes. Comparing results from the person's first and second languages would also be ideal to clarify this issue.

Second, the lack of significant interactions between prime type and language group for L1 and L2 bilinguals was surprising, given that some of these interactions yield significant differences when the data was analysed simply according to native language. This could suggest similarities exist in lexical organization based on being bilingual, rather than on the native language. Alternatively, as suggested by Vaid & Meuter (2017) it could indicate that monolinguals do differ from bilinguals regardless of whether the latter are tested in their first or second languages. Further work could try to test L1 English bilinguals in their second language, and L2 English Bilinguals in their first in order to see if they produce similar results.

In addition to unexpected findings, there are also unresolved questions, such as those of Hypothesis 3. There are avenues to explore in order to examine the hypothesis further, such as improving the complete language profiles of participants, which could help to understand the effectiveness of the online paradigm. To avoid the limitations of this project, future work on the subject should examine performance proficiency and self-reported proficiency via slider questions (e.g., 'Rate your level of Bilingualism'), or ranked choice questions (e.g., 'Rank your languages in terms of proficiency'), with possible translations of questions and instructions if there is uncertainty on the part of potential participants.

Conclusion

In summary, the data for the priming method gathered online appears to be sensitive enough to properly record semantic priming effects. This means that word association and lexical decision tasks can both be employed to study connections in the mental lexicon going forward. That these effects were accurately recorded for both monolinguals and bilinguals indicate that the paradigm can be used to measure the bilingual lexicon in future.

Additionally, the research findings appear to support the position that distinctions between syntagmatic and paradigmatic primes are not as pronounced as previously assumed, though some preference does remain depending on an L1 or L2 speakers' proficiency. However, to conclude whether this preference is a result of different background language experiences, further refining of the paradigm is required. The research also reveals the challenges in formulating complete linguistic profiles for individuals, as the diversity of learning styles, self-reported proficiency, daily usage, and competing processing strategies must be accounted for when creating measurable tasks. Yet, it is in learning to balance all these factors that we can better understand how lexical knowledge is stored and retrieved.

References

Auer, E. T. & Bernstein, L. E. (2008). Estimating When and How Words Are Acquired: A Natural Experiment on the Development of the Mental Lexicon. *Journal of Speech, Language, and Hearing Research*, *51*(3), 750–758. https://doi.org/10.1044/1092-4388(2008/053)

Borodkin, K., Livny, A., Kushnir, T., Tsarfaty, G., Maliniak, O., & Faust, M. (2021). Linking L2 proficiency and patterns of functional connectivity during L1 word retrieval. *Brain and Language*, *216*, 104931–104931. https://doi.org/10.1016/j.bandl.2021.104931

Burgess, C., & Simpson, G. B. (1988). Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language*, *33*(1), 86–103. https://doi.org/10.1016/0093-934x(88)90056-9

Chiu, W., & Lu, K. (2015). Paradigmatic relations and syntagmatic relations: How are they related? *Proceedings of the Association for Information Science and Technology*, *52*(1), 1–4. https://doi.org/10.1002/pra2.2015.1450520100122

Clark, L. (2012). *Investigating the syntagmatic-paradigmatic shift in second language speaking adults* (Order No. 1520179). Available from ProQuest Dissertations & Theses Global. (1143232755). Retrieved from https://proxy.library.brocku.ca/login?url=https://www-proquest-com.proxy.library.brocku.ca/dissertations-theses/investigating-syntagmatic-paradigmatic-shift/docview/1143232755/se-2?accountid=9744

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. Journal of *Verbal Learning and Verbal Behavior*, *8*, 240–247.

Collins, A. M., & Quillian, M. R. (1970). Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior*, *9*(4), 432–438

de Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical Decision and Word Naming in Bilinguals: Language Effects and Task Effects. *Journal of Memory and Language*, *47*(1), 91–124. https://doi.org/10.1006/jmla.2001.2840

Dóczi, B. (2020). An Overview of Conceptual Models and Theories of Lexical Representations in the Mental Lexicon. In S. A. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 46–65). essay, New York, NY.

Ervin, S. M. (1961). Changes with Age in the Verbal Determinants of Word-Association. *The American Journal of Psychology*, *74*(3), 361–372. https://doi.org/10.2307/1419742

Eddington, C. M., & Tokowicz, N. (2012). Examining English–German translation ambiguity using primed translation recognition. *Bilingualism: Language and Cognition*, *16*(2), 442–457. https://doi.org/10.1017/s1366728912000387

Fitzpatrick, T., & Izura, C. (2011). Word Association in L1 and L2. *Studies in Second Language Acquisition*, *33*(3), 373–398. https://doi.org/10.1017/s0272263111000027

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory & Cognition*, *11*(6), 565–574. https://doi.org/10.3758/bf03198281

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Horiba, Y. (2012). Word Knowledge and Its Relation to Text Comprehension: A Comparative Study of Chinese- and Korean-Speaking L2 Learners and L1 Speakers of Japanese. *The Modern Language Journal (Boulder, Colo.)*, *96*(1), 108–121. https://doi.org/10.1111/j.1540-4781.2012.01280.x

Kealtey, C., & de Gelder, B. (1992). The bilingual primed lexical decision task: cross-language priming disappears with speeded responses. *European Journal of Cognitive Psychology*, *4*(4), 273-292.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory & Language*, *33*, 149–174

Lee, Y., Jang, E., & Choi, W. (2018). L2-L1 Translation Priming Effects in a Lexical Decision Task: Evidence From Low Proficient Korean-English Bilinguals. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00267

Libben, G. (2022). Lexical representations in language processing. *Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.398

Libben, G., & de Almeida (2000). Is there a morphological parser? In S. Bendjaballah, W. U. Dressler, O. E. Pfeiffer & M. D. Voeikova (eds.), *Morphology 2000*, pp. 213-225. Amsterdam: John Benjamins, 2002.

Lucy, J. A. (2005). Through the Window of Language: Assessing the Influence of Language Diversity on Thought. *Theoria*, *54*, 299–309.

Malt, B. C., Li, P., Pavlenko, A., Zhu, H., & Ameel, E. (2015). Bidirectional lexical interaction in late immersed Mandarin-English bilinguals. *Journal of Memory and Language*, *82*, 86–104. https://doi.org/10.1016/j.jml.2015.03.001

Marian, V. & Kaushanskaya, M. (2007). Cross-linguistic transfer and borrowing in bilinguals. *Applied Psycholinguistics, 28*(2), 369-390. doi:10.1017/S014271640707018X

McNamara, T. P. (2005). *Semantic priming perspectives from memory and word recognition* . Psychology Press. ProQuest Ebook Central https://ebookcentral-proquest-com.proxy.library.brocku.ca

Mohsen, M. A. (2016). Effects of help options in a multimedia listening environment on L2 vocabulary acquisition. *Computer Assisted Language Learning*, *29*(7), 1220–1237. https://doi.org/10.1080/09588221.2016.1210645

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1003–1019. https://doi.org/10.1037/0278-7393.15.6.1003

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin*, *84*(1), 93–116. https://doi.org/10.1037/0033-2909.84.1.93

Odlin, T. (2005). Crosslinguistic Influence and Conceptual Transfer: What are the Concepts? *Annual Review of Applied Linguistics*, *25*, 3-25. doi:10.1017/S0267190505000012

Pan, X., & Jared, D. (2021). Effects of Chinese word structure on object perception in Chinese–English bilinguals: Evidence from an ERP visual oddball paradigm. *Bilingualism: Language and Cognition*, *24*(1), 111-123.

Pavlenko, A. (2009). Conceptual Representation in the Bilingual Lexicon and Second Language Vocabulary Learning. In *The bilingual mental lexicon: interdisciplinary approaches* (pp. 99–124). essay, Multilingual Matters.

Pavlenko, A. (2014). *The bilingual mind: and what it tells us about language and thought*. Cambridge University Press.

Runnqvist, E., Strijkers, K., & Costa, A. (2019). Error-based learning and lexical competition in word production: Evidence from multilingual naming. *PLOS ONE*, *14*(3). https://doi.org/10.1371/journal.pone.0213765

Sabourin, L., Brien, C., & Burkholder, M. (2014). The effect of age of L2 acquisition on the organization of the bilingual lexicon: Evidence from masked priming. *Bilingualism: Language and Cognition*, *17*(3), 542-555. doi:10.1017/S1366728913000643

Schwartz, A. I., & Arêas da Luz Fontes, A. B. (2008). Cross-language mediated priming: Effects of context and lexical relationship. *Bilingualism: Language and Cognition*, *11*(1), 95–110. https://doi.org/10.1017/s1366728907003239

Segalowitz, S. J., & Zheng, X. (2009). An ERP study of category priming: Evidence of early lexical semantic access. *Biological Psychology*, *80*(1), 122–129. https://doi.org/10.1016/j.biopsycho.2008.04.009

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, *11*(2), 245–260. https://doi.org/10.1017/s1366728908003404

Singh, L. (2014). One World, Two Languages: Cross-Language Semantic Priming in Bilingual Toddlers. *Child Development*, *85*(2), 755–766. https://doi.org/10.1111/cdev.12133

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 1*, 214–241.

Struck, J., & Jiang, N. (2021). Language switch costs in a lexical decision task: Symmetry and cognitive correlates. *Second Language Research*, 026765832199874. https://doi.org/10.1177/0267658321998740

Teske, K. (2017). Duolingo. *CALICO Journal*, *34*(3), 393–401. https://doi.org/10.1558/cj.32509

Vaid, J., & Meuter, R. F. I. (2017). Languages without borders: Reframing the study of the bilingual mental lexicon. In M. Libben, M. Goral, & G. Libben (Eds.), *Bilingualism: a framework for understanding the mental lexicon* (pp. 8–39). John Benjamins Publishing Company.

Vivas, J., Kogan, B., Romanelli, S., Lizarralde, F., & Corda, L. (2020). A cross-linguistic comparison of Spanish and English semantic norms: Looking at core features. *Applied Psycholinguistics*, *41*(2), 285–297. https://doi.org/10.1017/s0142716419000523

Yanguas, I. (2012). Task-based Oral Computer-mediated Communication and L2 Vocabulary Acquisition. *CALICO Journal*, *29*(3), 507–531.

Zareva, A. (2007). Structure of the second language mental lexicon: how does it compare to native speakers' lexical organization? *Second Language Research*, *23*(2), 123–153. https://doi.org/10.1177/0267658307076543

Zhang, D., & Koda, K. (2017). Assessing L2 vocabulary depth with word associates format tests: issues, findings, and suggestions. *Asian-Pacific Journal of Second and Foreign Language Education*, *2*(1). https://doi.org/10.1186/s40862-017-0024-0

APPENDICES

APPENDIX A

Table 6
Complete List of Target Words

Target Word	Syntagmatic Prime	Paradigmatic Prime	Clang Prime	Unrelated Prime
ANCHOR	marine	boat	banker	tame
APPLE	cider	pear	grapple	tomb
AXE	grind	sword	lacks	libel
BEACH	sandy	park	leech	lye
BELL	ringing	chime	spell	scale
BELT	leather	shirt	melt	collect
BOOK	read	magazine	hook	race
BREAD	butter	toast	thread	basil
BRIDGE	suspension	overpass	fridge	common
CAKE	birthday	pie	flake	disco
CANDY	sweet	cookie	dandy	effect
CANE	sugar	stick	stain	type
CAR	crash	motorcycle	far	pardon
CARPET	dusty	rug	tarpit	beast
CAT	litter	fox	mat	plate
CHAIN	supply	rope	main	alms
CHERRY	blossom	raspberry	scary	tanner
CHICKEN	grilled	turkey	stricken	hall
CHOIR	sing	vocalists	friar	birch
CHURCH	pastor	mosque	lurch	jogger
CITY	limits	village	pity	belief
CLOCK	time	sundial	stock	torrent
CLOUD	data	fog	proud	tight
CLOVER	lucky	thistle	rover	heed
CLOWN	circus	jester	gown	grid
COAL	black	oil	mole	interest
COAT	warm	jacket	goat	science
COFFEE	bean	tea	toffy	tractor
COIN	copper	money	join	chair
COW	dairy	cattle	now	dowel

CURTAIN	shower	window	uncertain	vendor
DANCER	ballet	musician	answer	sum
DAY	night	month	ray	pike
DEER	antlers	elk	clear	allure
DINNER	food	lunch	winner	micron
DOCTOR	hospital	nurse	proctor	myth
DOG	furry	wolf	log	twist
DOOR	sliding	entrance	gore	bound
EAGLE	bald	falcon	legal	princess
ECHO	sound	whisper	gecko	coolant
EGG	scrambled	sausage	beg	howl
ELBOW	joint	knee	jumbo	callous
FACE	mask	back	case	jump
FARM	livestock	barn	charm	vision
FIELD	track	meadow	healed	cause
FINGER	ring	toe	linger	north
FIRE	smoke	flame	tire	round
FLAG	country	banner	hag	arrival
FLOOR	dance	ceiling	score	mild
FLOWER	garden	grass	power	dike
FLY	wing	bee	try	bear
FOOT	walk	meter	soot	toucan
FRIEND	best	enemy	tend	era
GARAGE	mechanic	carpark	mirage	poor
GEAR	gadget	cog	sheer	pelt
GOLD	glitter	silver	mould	scribe
GOOSE	canada	swan	moose	colon
GYM	fitness	school	skim	depth
HAIR	comb	skin	fair	sire
HAMMER	tool	chainsaw	grammar	goalie
HARBOR	ships	airport	arbor	dollop
HEAD	brain	tail	bed	cassia
HEART	blood	lung	cart	trade
HILL	steep	cliff	spill	picket
HOSE	spray	rake	compose	shorten
HOUSE	haunted	apartment	louse	baker
IGUANA	green	crocodile	sauna	geode
JAIL	inmate	prison	shale	item
JAM	jar	jelly	scam	event

JAZZ	saxophone	hip-hop	has	carrion
JOB	menial	career	rob	impasse
JOURNAL	publish	diary	kernel	anger
JUG	gallon	bottle	hug	moist
JUICE	fruit	nectar	truce	stapler
KETTLE	boil	cauldron	settle	typhoon
KEY	lock	password	plea	olive
KING	crown	prince	sing	argue
KITE	windy	sail	might	tantrum
KITTEN	cute	puppy	mitten	handle
KNIFE	stab	fork	life	cycle
KNIGHT	armor	squire	right	peat
LADY	gaga	lord	shady	nuclear
LAKE	fishing	pond	take	aardvark
LAW	enforce	rule	jaw	teflon
LEEK	cooking	challot	meek	caution
LENS	camera	glasses	pens	expel
LIBRARY	quiet	bookstore	bribery	chemical
LIME	sour	orange	thyme	quell
LIVER	disease	kidney	quiver	matrix
MAIL	envelope	post	trail	balance
MAPLE	syrup	oak	staple	lantern
MEAL	eat	snack	teal	plasma
MEDAL	compete	award	pedal	trite
MOTOR	oil	piston	voter	span
MOUNTAIN	climb	valley	fountain	agree
MOUSE	cheese	rabbit	spouse	rant
MUSTARD	dijon	ketchup	custard	immune
NEON	glow	pastel	freon	tumble
NOODLE	soup	pasta	strudel	phyla
NOSE	smell	ear	pose	titan
NUT	crack	bean	hut	equip
OAR	canoe	paddle	more	height
OASIS	desert	sanctuary	basis	numeral
OCEAN	waves	sea	motion	timber
ONION	tears	garlic	bunion	born
PAIL	shovel	bucket	quail	fallow
PANTS	wear	shorts	chants	sultry
PEARL	oyster	jewel	squirrel	sickle

PEPPER	spicy	carrot	leper	film
PHONE	call	telegraph	scone	current
PIG	roast	hog	dig	breath
POT	cook	pan	thought	scan
POUND	british	gram	found	horrid
PUZZLE	play	game	muzzle	middle
QUEEN	royalty	empress	scene	messy
RAIN	heavy	hail	vain	index
RICE	fried	wheat	vice	fasten
RIVER	flow	stream	shiver	calorie
ROCK	roll	boulder	sock	loyal
ROOT	soil	branch	loot	tigress
ROSE	fragrant	lily	clothes	giant
RUBY	red	diamond	newbie	chowder
SAILOR	navy	pilot	tailor	cult
SCHOOL	learn	college	mule	pectin
SHOE	lace	sandal	grew	yearn
SIDEWALK	concrete	street	squawk	felt
SILK	tie	cotton	milk	portion
SKATE	ice	ski	fate	lemon
SNAIL	slimy	slug	hail	marrow
SNAKE	slither	lizard	fake	checkers
SNOW	cold	sleet	grow	target
SOAP	clean	shampoo	hope	dealer
SPARROW	nest	robin	narrow	yield
SPIDER	web	scorpion	glider	courage
SUMMER	hot	autumn	hummer	wheeze
MON	shine	moon	done	answer
TAPE	duct	glue	shape	lesson
THUNDER	storm	lightning	blunder	couched
TOOTH	dentist	eye	couth	amuse
TOWEL	bath	blanket	scowl	blow
TRAIN	rail	bus	gain	swell
TREE	leaf	bush	free	joy
TRUCK	driver	van	luck	shallow
TUBA	loud	trumpet	scuba	attack
TUNA	canned	salmon	lacuna	ivory
ULCER	stomach	tumor	pulsar	paltry
UMBRELLA	protect	parasol	patella	moody

VANILLA	flavor	chocolate	chinchilla	keg
VILLAGE	rural	hamlet	spillage	façade
VIOLIN	concert	cello	chagrin	beef
VIRUS	sick	bacteria	iris	brawl
WALL	brick	roof	call	league
WATER	drink	earth	potter	earnest
WHEAT	flour	barley	feat	hoax
WINE	ferment	beer	spine	hound
WINTER	freeze	spring	printer	income
WOOL	sheep	fleece	pull	cumin
YARN	knitting	fabric	darn	tavern
YOGA	mat	meditation	toga	cancel
ZEBRA	stripes	horse	libra	elegant

Table 7

Complete List of Target Non-Words

Target Word	**Clang Prime**	**Unrelated Prime**
AGRIL	nostril	soca
ALK	chalk	well
ANGEST	longest	seem
APLEE	sharply	cousin
ARLIK	bootlick	service
ATHY	wealthy	lead
AYRESS	terrace	revenge
BALOR	valor	real
BAMP	cramp	version
BANOT	unknot	insight
BELUE	venue	flex
BIRT	hurt	flesh
BLEER	fear	rare
BLUTE	flute	unit
BOSESS	possess	pitch
BOST	glossed	marsh
BOTTING	rotting	sniff

BRELT	dealt	air
CHEMMY	tummy	moral
CLARD	hard	daughter
CLEATH	sheath	unique
COTEE	goatee	request
COUDAY	holiday	joke
CRIX	tricks	mold
CROOM	loom	ditch
CROTH	sloth	colon
DIMIN	ermine	fix
DIPPANT	flippant	waist
DISSET	corset	pudding
DOIL	toil	retire
DOPACT	contact	tongue
DREE	pea	honor
DREEL	seal	host
DROOM	broom	excuse
DROSE	prose	swipe
EMPUSE	abuse	mild
ENCE	hence	ripe
ESHUR	gesture	asset
FAMBER	timbre	bracket
FAZZEL	dazzle	attract
FLANE	slain	love
FORN	torn	freckle
FREST	chest	nail
FUVVER	lover	sow
GINE	mine	warrant
GISS	hiss	peace
GLALE	trail	battle
GLEER	fear	think
GOX	socks	flourish
GOYETO	falsetto	panel
GREAM	dream	ribbon
GRUN	spun	fade
GUCH	much	hawk
GUFFIN	muffin	healthy
GUNG	hung	ladder

GUTTLE	scuttle	toll
HANE	vane	exempt
HERM	worm	achieve
HONWIN	penguin	blade
HOREN	siren	defend
HUBIER	stubbier	assault
HUDEN	madden	constant
HUNGE	lunge	clinic
IME	rhyme	captain
INVAKE	forsake	killer
JEG	leg	lump
JERL	earl	kinship
JILA	villa	ticket
JIRE	pyre	office
JUR	fur	troop
KELL	tell	fever
KIG	fig	theory
KINAP	clap	clerk
KIRE	spire	yard
KLUST	trust	coffin
KNAST	mast	spray
KONER	loner	crop
KOSS	moss	hate
KOSTOVY	anchovy	volume
LABE	babe	immune
LAMACH	damage	debt
LAWI	kiwi	robot
LETTISM	autism	outlook
LINACT	react	pillow
LIRRER	mirror	good
LOBESS	obsess	insure
LOMBER	somber	dome
LOUN	gown	stairs
LUMIT	suet	enhance
MANG	bang	faithful
MAPE	cape	sit
MASHA	fuscia	bait
MASIDE	decide	letter

MEMPER	whimper	team
MIMPLE	pimple	practice
MOUT	trout	tablet
MUGESS	process	tent
MUN	bun	welcome
MURT	dirt	utter
NENT	sent	displace
NILT	tilt	branch
NOPILA	gorilla	buy
NOTH	goth	surgeon
NUNK	monk	guard
OME	home	neutral
OZZED	caused	trace
PAMP	camp	feign
PANBIN	carbon	vat
PAPPLE	dapple	cry
PAQUIL	tranquil	formal
PAWT	squat	unrest
PEDISH	radish	tactic
PINDER	gender	rub
PLIVE	jive	slam
PRENT	tent	crevice
RAL	small	inflate
RAMPER	jumper	first
RAPISH	sheepish	blame
RAPPON	tampon	soft
REJAKE	mistake	treaty
REPICE	hospice	ignite
RETICT	predict	iron
RINNEL	tunnel	depress
RINT	stint	chimney
RITMER	glimmer	mislead
RONT	gaunt	square
RULL	skull	cruelty
SATOTE	remote	quota
SAWN	fawn	gate
SEN	when	evoke
SERDER	herder	thick

SIRKY	quirky	allow
SKEM	them	crisis
SLEAD	mead	exam
SOTHY	frothy	fiction
SPAVE	grave	brown
SPUK	tuck	gaffe
STASS	grass	pension
STINER	whiner	organ
SUSTER	cluster	kidnap
SYLING	tiling	endure
TAND	land	obese
TEED	seed	grand
THOP	shop	orbit
TODER	fodder	physics
TREL	spell	slump
TRINE	brine	choice
TROSTY	frosty	baby
TRUSH	mush	pour
TWIL	quill	movie
TWOP	quip	help
TWOFT	loft	beam
VARROW	harrow	disclose
VIFFIN	puffin	rib
WAM	dam	legend
WERN	turn	tough
WIB	fib	symbol
WIFT	shift	pass
XECO	gecko	chapter
ZAYSA	mesa	carry

APPENDIX B

English Proficiency Questionnaire	
Press the number key associated with your answer.	
I can introduce myself and say what my job is.	1 = I cannot do this 2 = Poorly 3 = With Difficulty 4 = Reasonably Well 5 = I can do this fluently
I can tell the date and time.	
If someone calls and speaks English on the phone, I can take and communicate basic messages.	
I can provide directions to a location (for example, to a store).	
I can say what I plan to do during the weekend.	
I can describe a foreign city that I have visited.	
I can explain why a certain location is my favourite vacation spot by providing three reasons.	
I can refuse an invitation to go to somewhere (for example, to a restaurant) and provide reasons for my refusal.	
I can say what I did yesterday.	
I can start a conversation when I meet someone.	
I can return an item to a store and explain why I am returning that item.	
I can phone a doctor's office to make an appointment and explain my medical issue in general terms.	
I can praise or criticize a film, a play or a person.	
I can contrast two cities in the world on the basis of lifestyle and culture.	
At a restaurant, I can complain about a dish and express my displeasure.	
After a presentation, I can ask a question.	
I can formulate a suggestion at a formal meeting.	
If I do not share someone's opinion on a	

serious matter, I can argue and express my opinion.	
I can talk about the pros and cons of a particular situation.	
I can provide a detailed account of a speech given by the head of the organization for which I work.	

Language Learning Questionnaire

The following questions will go into greater detail about the languages you have learned, including English. If a question does not apply to your situation, please select the appropriate option.

1. Do you consider yourself to be bilingual or multilingual?
    a. Yes
    b. Somewhat
    c. No
2. At what age did you first learn a second language?
    a. Since Birth
    b. Before age 7
    c. Before age 15
    d. Before age 30
    e. Before age 60
    f. I have not learned a second language
3. How long have you been speaking a second language?
    a. I do not speak a second language
    b. Less than a year
    c. 1-2 years
    d. 3-5 years
    e. 5-10 years
    f. 10+ years
4. How did you first learn a second language? (Select all that apply)
    a. I don't speak a second language
    b. From family
    c. From classes
    d. At work
    e. From printed media
    f. From broadcast media
    g. From social media
    h. Other (please specify)
5. If you answered question 4 with any "media" option (e, f, or g), which of the following did you use to learn a second language? (Select all that apply)
    a. Not Applicable
    b. Music
    c. Television & Movies
    d. Messaging apps (e.g., WhatsApp, Messenger, etc.)

   e. Language Learning apps (e.g., Babbel, Duolingo, etc.)
   f. Books & Comics
   g. News
   h. Others (please specify)
6. Where do you use a second language most often?
   a. I don't speak a second language
   b. I don't use my second language
   c. At home
   d. At work
   e. At school
   f. At a place of worship
   g. In public
   h. Other (Please specify)
7. With whom do you converse with in a second language?
   a. I don't speak a second language
   b. I don't use my second language
   c. With immediate family
   d. With extended family
   e. With friends
   f. With coworkers
   g. With employers
   h. Other (please specify)

8. How often do you use your Second Language?
   a. I don't speak a second language
   b. Never
   c. Less than once a month
   d. Once or twice a month
   e. Once or twice a week
   f. Several times per week
   g. Daily
   h. Other (please specify)

APPENDIX C

Table 8

Participants' Accuracy for Target Words in the Pilot Study

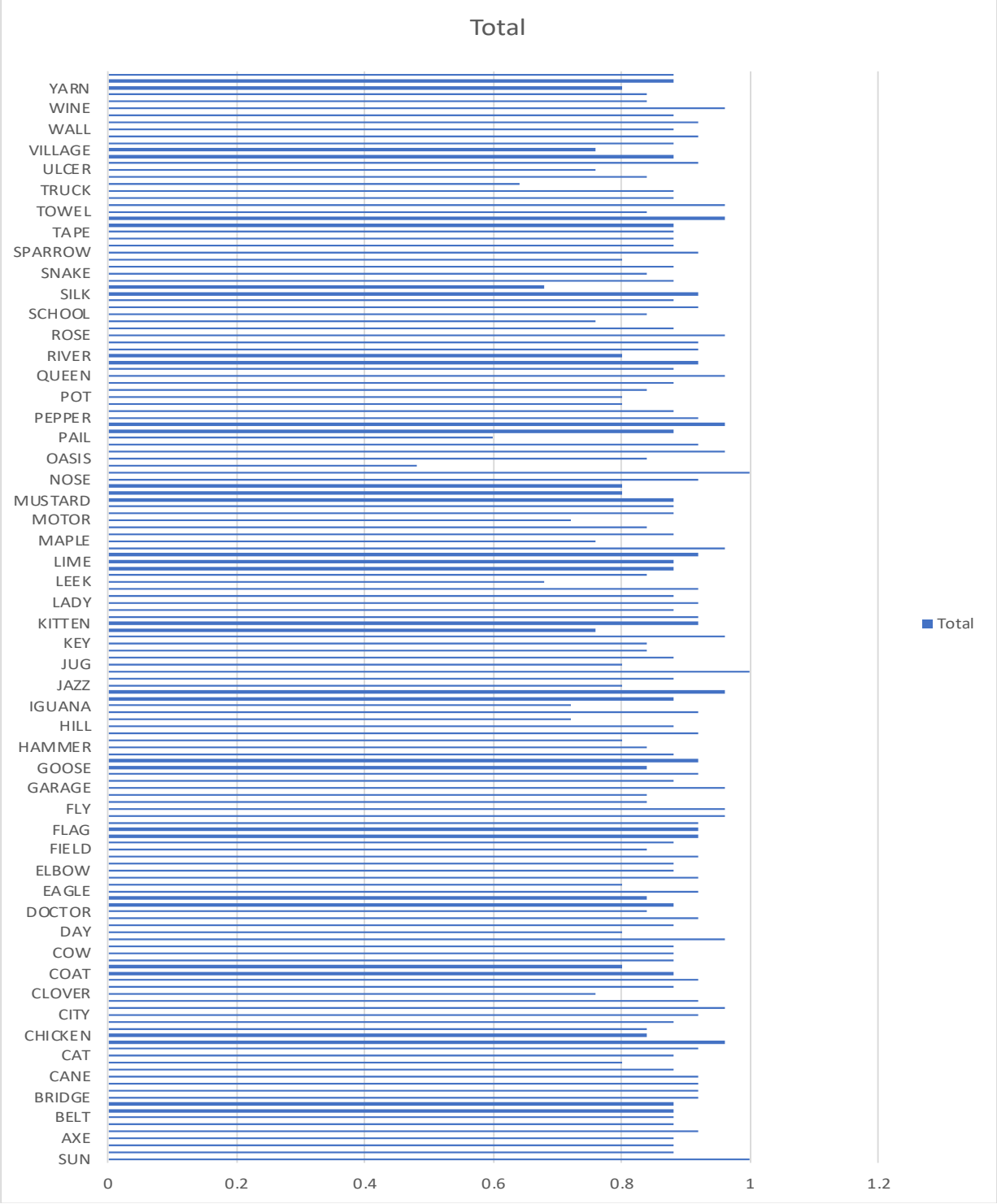


Table 7

Participants' Accuracy for Target Non-Words in the Pilot Study

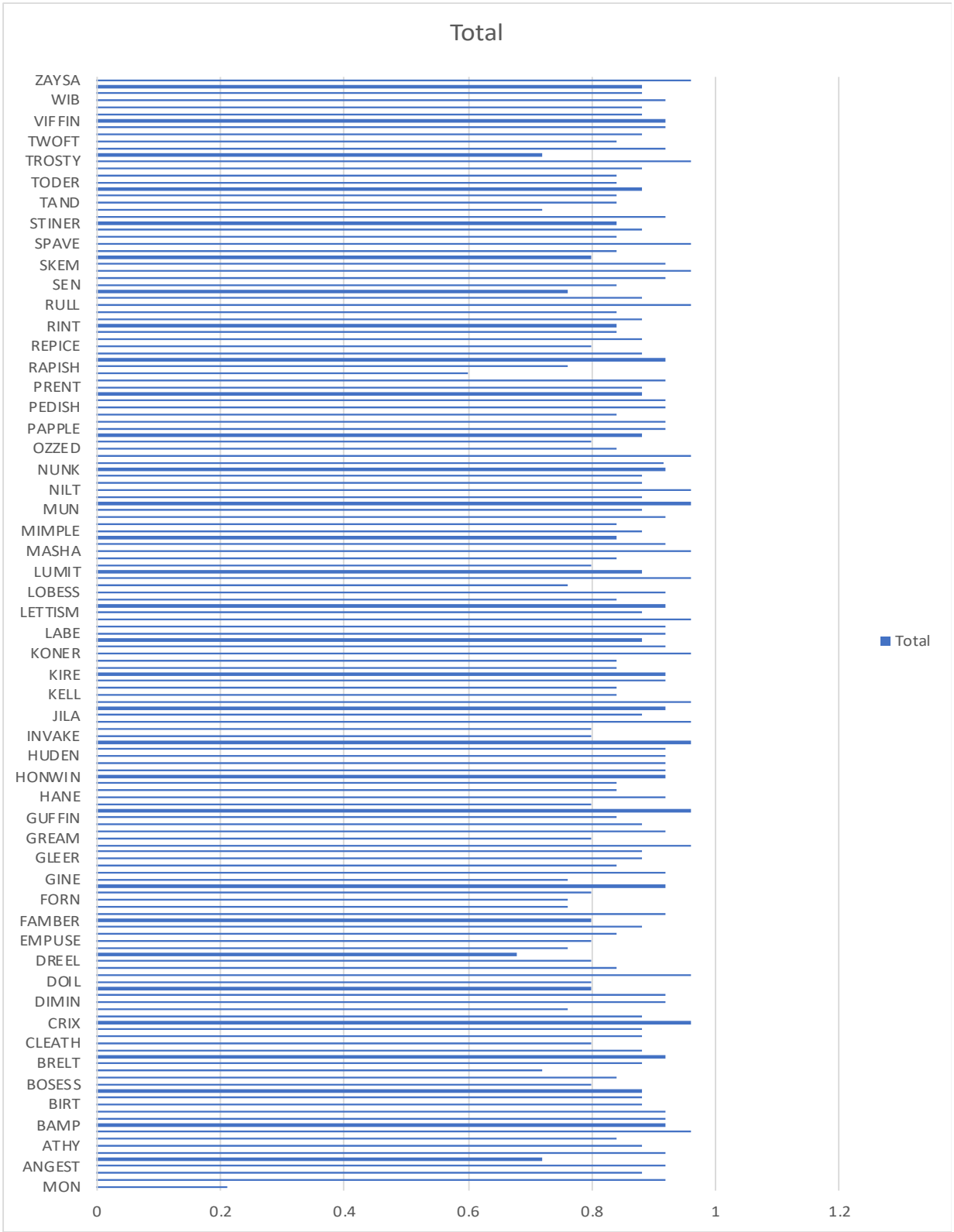


Table 10

Estimated Marginal Means Contrasts between Linguistic Groups

Contrast	Estimate	SE	df	Z ratio	P value
L1 Bilingual - L1 Monolingual	0.05	0.04	Inf	1.20	0.45
L1 Bilingual - L2 Bilingual	0.18	0.04	Inf	4.72	<.0001
L1 Monolingual - L2 Bilingual	0.13	0.04	Inf	3.05	0.0066

Table 11

Estimated Marginal Means between Prime Types' Facilitation Effects

Contrast	estimate	SE	df	Z ratio	P value
UNRELATED - CLANG	0.009000	0.00385	Inf	2.336	0.0899
UNRELATED - PARADIGMATIC	0.021182	0.00388	Inf	5.461	<.0001
UNRELATED – SYNTAGMATIC	0.022162	0.00391	Inf	5.667	<.0001
CLANG - PARADIGMATIC	0.012182	0.00385	Inf	3.161	0.0085
CLANG - SYNTAGMATIC	0.013161	0.00389	Inf	3.385	0.0040
PARADIGMATIC – SYNTAGMATIC	0.000979	0.00392	Inf	0.250	0.9945

Table 12

LMER Fixed Effects for Mean Reaction Time between Participants' Native Language

Variable	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.18	0.07	315.95	109.35	< 0.001 ***
Prime CLANG	0.00	0.00	27825.37	0.03	0.98
Prime PARADIGMATIC	-0.01	0.00	27792.30	2.70	0.007 **
Prime SYNTAGMATIC	-0.02	0.00	27802.37	3.82	0.001 ***
Native Language (English L2)	-0.18	0.03	211.10	5.29	0.001 ***
Previous Reaction Time	0.00	0.00	27799.50	1.05	0.29
Trial Number	0.00	0.00	27800.32	0.73	0.47
Age	0.00	0.00	202.96	1.67	0.1 .
Gender (Woman)	0.06	0.03	202.97	2.14	0.03 .
Target Frequency (Log)	-0.03	0.00	158.97	11.48	< 0.001 ***
Education (College or Technical School)	-0.21	0.05	202.69	3.72	< 0.001 ***
Education (High School)	-0.21	0.05	202.72	4.51	< 0.001 ***
Education (Less than High School)	-0.21	0.21	202.64	0.97	0.33
Education (Master's/Doctorate)	-0.01	0.04	203.00	0.24	0.81
Target Length	0.00	0.00	153.77	0.07	0.95
Clang x English L2	-0.02	0.01	27770.59	-2.31	0.02 .
Paradigmatic x English L2	-0.02	0.01	27771.26	-2.51	0.01 *
Syntagmatic x English L2	-0.01	0.01	27768.50	-1.48	0.14

Table 13

LMER Fixed Effects on Mean Reaction Time according to Participants' Linguistic Status (PLS)

Value	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.15	0.07	313.03	107.93	< 0.001 ***
Prime CLANG	0.00	0.01	27809.44	-0.58	0.56
Prime PARADIGMATIC	-0.01	0.01	27787.12	-2.76	< 0.01 **
Prime SYNTAGMATIC	-0.02	0.01	27792.13	-3.69	< 0.001 ***
PLS L1 Monolingual	-0.05	0.04	209.83	-1.25	0.21
PLS L2 Bilingual	-0.17	0.04	209.83	-4.46	< 0.001***
Previous Reaction Time	0.00	0.00	27796.06	1.06	0.29
Trial number	0.00	0.00	27797.09	-0.71	0.47
Age	0.00	0.00	201.98	-1.14	0.26
Gender (Woman)	0.07	0.03	201.98	2.17	0.03 .
Target Frequency (log)	-0.03	0.00	158.97	-11.48	< 0.001 ***
Education (College or Technical School)	-0.20	0.06	201.71	-3.59	< 0.001 ***
Education (High School)	-0.20	0.05	201.74	-4.15	< 0.001 ***
Education (Less than High School)	-0.40	0.22	201.65	-1.82	0.07 .
Education (Master's/Doctorate)	-0.01	0.04	201.99	-0.19	0.85
Target Length	0.00	0.00	153.77	0.07	0.95
Clang x L1 Monolingual	0.01	0.01	27773.05	0.72	0.47
Paradigmatic x L1 Monolingual	0.00	0.01	27774.29	0.21	0.84
Syntagmatic x L1 Monolingual	0.00	0.01	27770.57	0.07	0.94
Clang x L2 Bilingual	-0.02	0.01	27776.80	-1.80	0.07 .
Paradigmatic x L2 Bilingual	-0.01	0.01	27785.14	-1.59	0.11
Syntagmatic x L2 Bilingual	0.00	0.01	27774.31	-0.43	0.67

Table 14

LMER Fixed Effects of Reaction Times in L2 Bilinguals

Value	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.63	0.33	50.28	22.95	< 0.001 ***
Prime CLANG	-0.02	0.01	7338.39	-2.43	0.02 .
Prime PARADIGMATIC	-0.03	0.01	7331.91	-4.04	< 0.001 ***
Prime SYNTAGMATIC	-0.02	0.01	7320.45	-3.44	< 0.001 ***
Participant Average LDT accuracy	-0.92	0.36	49.17	-2.58	< 0.01 *
Age	0.01	0.00	48.89	3.04	< 0.01 **
Gender (Woman)	-0.03	0.04	48.95	-0.87	0.39
Target Frequency (log)	-0.04	0.00	161.73	-11.20	< 0.001 ***

Table 15

LMER Fixed Effects of Reaction Times in L1 Bilinguals

Value	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	8.27	0.27	104.09	30.45	< 0.001 ***
Prime CLANG	0.00	0.01	13306.97	-0.53	0.60
Prime PARADIGMATIC	-0.01	0.01	13281.30	-2.63	< 0.01 **
Prime SYNTAGMATIC	-0.02	0.01	13290.71	-3.48	< 0.001 ***
Participant Average LDT accuracy	-1.27	0.29	102.03	-4.40	< 0.001 ***
Age	0.00	0.00	102.06	-1.76	0.08 .
Gender (Woman)	0.08	0.04	102.02	1.74	0.08 .
Target Frequency (log)	-0.03	0.00	160.08	-9.50	< 0.001 ***

Table 16

LMER Fixed Effects of Reaction Times in L1 Monolinguals

Value	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	8.94	0.43	48.40	20.79	< 0.001 ***
Prime CLANG	0.00	0.01	7003.47	0.41	0.68
Prime PARADIGMATIC	-0.01	0.01	6979.56	-1.78	0.07 .
Prime SYNTAGMATIC	-0.02	0.01	6991.14	-2.72	0.00662 **
Participant Average LDT accuracy	-2.15	0.48	48.01	-4.44	< 0.001 ***
Age	0.00	0.00	47.95	-0.22	0.82
Gender (Woman)	0.04	0.07	47.96	0.53	0.60
Target Frequency (log)	-0.03	0.00	157.95	-10.12	< 0.001 ***

Table 17

LMER Fixed Effects of Reaction Times in All L1 Speakers

Value	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	8.45	0.21	158.28	40.05	< 0.001 ***
Prime CLANG	0.00	0.00	20410.22	-0.24	0.81
Prime PARADIGMATIC	-0.01	0.00	20369.35	-3.21	0.001 **
Prime SYNTAGMATIC	-0.02	0.00	20385.84	-4.39	< 0.001 ***
Participant Average LDT accuracy	-1.51	0.23	154.11	-6.49	< 0.001 ***
Age	0.00	0.00	154.00	-1.93	0.06 .
Gender (Woman)	0.08	0.04	154.01	2.06	0.04 .
Target Frequency (log)	-0.03	0.00	158.80	-10.76	< 0.001 ***