Investigating the Portuguese-English Bilingual Mental Lexicon: Crosslinguistic Orthographic and Phonological Overlap in Cognates and False Friends

Leonardo Alves-Soares

Thesis submitted to the

University of Ottawa

In partial fulfillment of the requirements

for the Doctorate in Philosophy degree in Linguistics

Department of Linguistics

Faculty of Arts

University of Ottawa

Abstract

This dissertation investigates how cognates are organized in the bilingual mental lexicon and examines whether orthography in one language, via phonological representations, influences the processing of cognates and false friends in the other language. In light of the framework of two well-known models of bilingual visual word recognition, the Bilingual Interactive Activation (BIA) and the Bilingual Interactive Activation Plus (BIA+), the premise is that there is activation from orthography to phonology across a bilingual's two languages and that this activation is modulated by the degree of orthographic and phonological code overlap.

Two objective metrics were used to assess crosslinguistic similarity of Portuguese-English cognates and false friends that were selected for a cross-language lexical decision task with masked priming. Dynamic time warping (DTW), an algorithm that was originally conceived to compare different speech patterns in automatic speech recognition and to measure acoustic similarity between two time-dependent sequences, was used to compute crosslinguistic phonological similarity. The Normalized Levenshtein Distance (NLD), an algorithm that calculates the minimum number of single-character insertions, deletions or substitutions required to change one word into another and normalizes the result by their lengths, was used to compute crosslinguistic orthographic similarity. Portuguese-English bilinguals who acquired their second language after reaching puberty, and English functional monolinguals who grew up speaking primarily English were recruited to participate in the experimental task.

Based on collected reaction time and accuracy data, mixed-effects models analyses are used to estimate the individual effects of crosslinguistic orthographic, phonological and semantic similarity and the role each of them, along with English proficiency, word frequency and length play in the organization of the Portuguese-English bilingual mental lexicon.

Keywords: bilingual mental lexicon, Portuguese, English, Portuguese-English bilinguals, English functional monolinguals, cognates, false friends, crosslinguistic overlap, crosslinguistic similarity, semantic, lexical decision task, LDT, masked priming, spreading activation, RHM, BIA, BIA+, Multilink, orthographic representations, phonological representations, dynamic time warping, DTW, Normalized Levenshtein Distance, NLD, mixed-effects models, `lmer`, `glmer`, reaction time, accuracy

Acknowledgements

If there is one thing 2020 will be remembered for is that it has been a very tough year for everyone due to COVID-19. These are strange times. Besides social isolation, physical distancing and having to wear masks in order to protect ourselves and others, there is much uncertainty about the future all around. Finishing a dissertation in 2020 makes acknowledging those that stuck by me and helped me through this journey even more important.

First and foremost, I would like to express my sincerest and perpetual gratitude to my PhD supervisor, Juana Muñoz-Liceras, who took a genuine interest in my research and never gave up on me even at the times when I felt there was no point moving ahead. Juana always encouraged me to keep my head up, shoulders back and to move ahead at full throttle gracefully, never looking back. Your wit, passion and tenacity, even when the politics of the world of academia get in the way, are an inspiration to me. I am extremely proud of having had the opportunity to work closely with you and consider it a privilege being among one of your doctoral students.

I would like to thank professor Marc Brunelle for opening my mind to the possibilities with mixed-effects models and helping me see a much brighter and more exciting world beyond ANOVAs. I will be forever grateful for all the time you took to get me started with R and mixed-effects models. It was not an easy road for me to navigate in the beginning, particularly when all I knew since my bachelor's in psychology was SPSS; however, you guided me in the right direction.

I also would like to thank all my committee members, Anahi Alba de la Fuente, Elena Valenzuela, Nikolay Slavkov and Ricardo Augusto de Souza, not only for taking the time to carefully read my work and provide me with invaluable feedback, but also for having made my dissertation defense feel more like a conversation among academics. It was certainly one of the hallmarks of this entire experience.

I owe my friend and fellow academic, Estela García-Alcaraz, an enormous thank you for all her assistance with reviewing my data analysis chapter, and, in particular, for going over my mixed-effects models with a fine-tooth comb. They say there is no better way to learn than to teach; I can say I have learned a great deal by imparting my knowledge of running mixed-effects models in R to you. Your challenging questions pushed me to delve deeper into my analyses and to literally go the extra mile.

I would like to recognize a few specific Faculty of Arts members who ensured my tenure at the University of Ottawa as a doctoral student was as rewarding as possible. They are Vice-Dean Frans De Bruyn, Rachel Fontaine-Azzi, Christine Bertrand, Suzanne Dalrymple and Roxanne Lacelle. In addition, I would like to express my appreciation to my colleagues in the English Intensive Program at the Official Languages and Bilingualism Institute at the University of Ottawa, who were willing to serve as functional monolingual participants in my experimental task.

I would like to thank my family in Brazil for supporting me to pursue my dreams in North America. And, finally, Pumba, thank you for always being there for me unconditionally along the way, through thick and thin. You have always been my biggest cheerleader, helping me keep my eyes on the prize and not lose track of a much brighter future. I would have never been able to accomplish all of this without you. And, to Caio Caiana, my little girl, thank you for being my furball, jumping on my desk when I was writing this dissertation, keeping me calm, and for showering me with lots of meows and purrs.

Table of Contents

List of Tables

List of Figures

List of Examples

List of Appendices

1. Introduction

For a Portuguese-English bilingual – or any other bilingual for that matter – riding the subway in New York City can be a truly remarkable experience. Among all the hustle and bustle, all the bright, flashing lights and all the announcements in English, the feeling of reading or hearing a Portuguese word in an English-speaking country can lead to moments of confusion and ambiguity. For example, seeing the English word *push* printed on access doors to stations, emergency exits or on the windows in the trains can automatically trigger the Portuguese word *puxe,* which sounds similar to the English word but has a completely different meaning in the language. *Puxe* in Portuguese means to exert force on someone (or something) so as to cause movement toward oneself (or itself), which in English is conveyed by the word *pull*. In other words, the pronunciation similarities between *puxe* and *push* are merely superficial. Similarly, the English word *platform* can automatically trigger the Portuguese word *plataforma*, except that these two words have the same meaning across both languages. Words like *puxe* and *push*, which share some phonemes, graphemes and yet have completely different meanings crosslinguistically, are often defined in the literature as *false friends*, or sometimes *false cognates.* Words like *platform* and *plataforma* are defined as *cognates*. They share many graphemes, phonemes and meaning crosslinguistically. Despite the fact that English is a Germanic language and Portuguese is a Romance language typologically, both languages have a considerable number of words whose etymological origin is Latin. As such, both languages also have considerable number of cognates, e.g., *computador* and *computer*, *piano* and *piano*, *televisão* and *television*, *robô* and *robot*, *zero* and *zero*, etc. English and Portuguese also possess a considerable number of false friends, e.g., *batom* 'lipstick' and *baton*, *chute* 'kick' and *chute*, *sapo* 'frog' and *sap*, *tampão* 'drain cover' and *tampon*, etc.

The extant psycholinguistic literature on the processing of cognates and false friends has helped elucidate the organization of the bilingual mental lexicon, particularly the mechanisms that trigger that sense of ambiguity that bilinguals can often experience when they read or hear a word in one of their languages that look and sound like a word in the other language. For example, in translation recognition and lexical decision tasks with masked priming, in which accuracy and reaction time are the dependent variables, highly fluent bilinguals show a significantly greater priming effect for cognates than for non-cognates in both L1-to-L2 and L2-to-L1 directions (Dunabeitia et al., 2010). In other words, orthographic, semantic and phonological similarity crosslinguistically seem to play a role in the organization of the mental lexicon. Thus, it is possible that crosslinguistically cognates may share a single mental representation. There is evidence suggesting that bilinguals are able to name cognates more accurately and faster than both non-cognates and false friends (Sunderman & Schwartz, 2008). Additionally, Jared and Szucs (2002) observed that French-English bilinguals were slower to name interlingual homographs (false friends with a great deal of crosslinguistic orthographic overlap with distinct phonological codes) than matched English controls in an English naming task. Similarly, Schwartz, Kroll and Diaz (2007) showed highly proficient English-Spanish bilinguals displayed greater naming latencies (or RTs) for cognates with high crosslinguistic orthographic overlap when their corresponding phonological codes were different crosslinguistically. This effect was even observed irrespective of the bilinguals' dominant language. Finally, there is evidence showing there is cross-language feed-forward activation from orthography to phonology, like during a visual word recognition task; cognates with a high degree of crosslinguistic orthographic overlap that sound rather different crosslinguistically are processed more slowly than cognates whose degree of orthographic and phonological code overlap is equal (Schwartz & Kroll, 2006).

Despite all of the evidence supporting the special status of cognates, many outstanding issues still remain. For example, most studies do not make a clear distinction among the various kinds of cognates, viz., true vs. false friends, semi-false vs. false friends, interlingual homophones vs. homographs, etc. Additionally, the effects of cross-language activation during the processing of cognates with varying degrees of phonological similarities (variations of how differently they sound crosslinguistically) has not been yet established for many language pairs, in particular English and Portuguese.

In light of this, the primary goal of this dissertation is to investigate how cognates are organized in the Portuguese-English bilingual mental lexicon and to understand whether orthography in Portuguese, via phonological representations, influences Portuguese-English bilinguals'[1] processing of cognates and false friends in their second language. The premise is that there is activation from orthography to phonology across a bilingual's two languages and that this activation is modulated by the degree of crosslinguistic orthographic and phonological code overlap. As such, this dissertation will further advance the understanding of the interplay between orthography and phonology in the organization of the bilingual mental lexicon.

Two groups of participants were recruited for a cross-language lexical decision task with masked priming, Portuguese-English bilinguals and English functional monolinguals, who served as controls. The stimuli selected for the experiment were comprised of Portuguese-English cognates and false friends. Their degree of crosslinguistic orthographic and phonological code overlap was determined through objective metrics such as the Normalized Levenshtein Distance and dynamic time warping. Two mixed-effects models were created to account participants' reaction times and accuracy.

This dissertation is organized as follows.

Chapter 2 offers a broad overview of early literature on the organization of the

[1] While in the extant literature there are multiple definitions of bilingualism, this study is primarily concerned with late sequential bilingualism, as defined by Montrul (n.d., p. 17). Specifically, Portuguese speakers who acquired English as their second language (L2) during or after puberty.

mental lexicon, from its origins on spreading activation and connectionist models, to current literature on the organization of the bilingual mental lexicon. A synopsis of the structural elements of five well-known models of bilingual word recognition, viz., the Revised Hierarchical Model (RHM), the Interactive Activation Model (IAM), the Bilingual Interactive Activation Model (BIA) and its derivatives, the Bilingual Interactive Activation Model Plus (BIA+) and Multilink, is also provided.

Chapter 3 explains the methodology employed in the study, from the process of participant recruitment and English proficiency testing to how all the stimuli were selected for the cross-language lexical decision task using objective metrics of crosslinguistic code overlap as well as how the experimental word lists were assembled.

Chapter 4 reports the results of the cross-language lexical decision task. It is divided into two sections, Preliminary Analyses and Mixed-Effects Models, with the former serving as groundwork for the latter. In Preliminary Analyses, each of the variables of interest are analyzed using conventional statistical analyses tools, such as t-tests, ANOVAs and correlations. In Mixed-Effects Models, a linear mixed-effect model (`lmer`), with reaction time as the dependent variable, and a generalized linear mixed-effect model (`glmer`), with accuracy as the dependent variable, are presented. Both mixed-effects analyses allowed for precise quantification of the individual fixed effects of crosslinguistic similarity, as well as their interactions with one another, which was one of the primary goals of this study.

Chapter 5 discusses the key findings from the experimental components of this study and formalizes the contribution to the existing research in light of the framework of two important models of the bilingual mental lexicon, the BIA and the BIA+.

Chapter 6 summarizes the study and offers the main conclusions as well as suggestions for further research.

2. Theoretical Background and State of the Art

2.1. Spreading Activation as the Foundation of the Mental Lexicon

The mental lexicon is theorized to be a mental dictionary that stores information pertaining to meaning, spelling and pronunciation of all the words a speaker has acquired throughout their life. One of the greatest challenges when proposing a model of the mental lexicon concerns the representation of concepts and form, specifically, how the conceptual representation of a word in memory (semantics) can activate graphemes (orthographic representations) and phonemes (phonological representations) and vice-versa. Collins and Quillian (1969), based upon spreading activation, proposed a model of lexicosemantic encoding and decoding in which they envisaged the mental dictionary as a network of interconnected nodes. In their view, the network nodes represent acquired linguistic knowledge and the pathways, or the connections/links between the nodes, represent their semantic associations. When a node in the network becomes activated, it automatically spreads (or percolates) its activation throughout its pathways iteratively (in a chain-like reaction) to other nodes with which it is semantically associated. For example, when the word *plant* is presented to an English speaker, the network nodes linked to the conceptual representation of the word in memory spread (or percolate) their activation to nodes linked to the conceptual representation of the word *flower*, which, in turn, spread their activation to nodes linked to the conceptual representation of the English word *rose*. Words in the mental lexicon are also thought to have a resting-level (baseline level) activation; words that are repeatedly activated together are posited to collectively retain a portion of their activation; thus, they are posited to be at a higher resting-level activation, which facilitates the speed with which activation spreads between them and to other associated words (Anderson, 1983a, 1983b, 2015; Samani & Sharifian, 1997).

Over the years the underpinnings of the mental lexicon were further advanced. For example, a hierarchical organization consisting of three layers of linguistic representation was proposed: a semantic or conceptual layer, which sits atop the hierarchy, an intermediate lexical layer, and a sublexical layer, which sits at the bottom of the hierarchy. In this framework, during the course of visual word recognition, activation is thought to spread iteratively, i.e., in a feed-forward fashion, within each level of linguistic representation as well as backwards from higher to lower level representations. In addition, the existence of a lexical selection system, which is responsible for ensuring that words in the mental lexicon that match a speaker's communicative intention are properly retrieved from semantic memory and produced is also introduced (Caramazza, 1997; Costa, La Heij, & Navarrete, 2006; Costa & Santesteban, 2004; Dell, 1986; Levelt, 1993, 2001; Levelt, Roelofs, & Meyer, 1999; Motley & Camden, 1985). Levelt (1993, 2001) argued that without a lexical selection system, spreading activation would create an insurmountable obstacle for communication, because a profusion of lexical items (or words) would be simultaneously active and compete with one another for production, which would make human communication a rather slow and error-ridden process. Levelt states that speakers *err* no more than once or twice every 1,000 words, in spite of the fact that they amass a mental lexicon of 50,000-100,000 words throughout their lifetime. Putting it simply, because human communication is both effective and fluid, a lexical selection system must be responsible for ensuring a speaker's intention is conveyed in a message. Levelt theorizes that the lexical selection system takes into account the amount of activation lexical and sublexical representations receive from semantic representations in order to ultimately select a lexical item with the highest level of activation, which in normal communication should match the speaker's communicative intention. Once selected, this lexical item is sent to subsequent linguistic processing or articulation.

2.2. The Bilingual Mental Lexicon

Highly proficient sequential bilinguals, without much apparent delay or effort, are able to switch between their L1 and L2 as well as map a specific word they read or hear to one of their two languages. This remarkable ability has been the subject of much investigation. Based on the enormous body of research accumulated over the last four decades, two theoretical issues have consistently remained under the microscope: (a) whether bilinguals' two languages are stored together in a single, integrated mental lexicon, expressed in the literature as the *shared-lexicon hypothesis*, and (b) whether the words from bilinguals' two languages are accessed simultaneously (in parallel) during word recognition, expressed in the literature as the *language non-selective hypothesis*. As a result, a number of models of bilingual word recognition with different core components, some leapfrogging earlier models, have been postulated over the years and tackle the shared lexicon hypothesis and the language non-selective hypothesis in diverse ways. The following is a synopsis of five well-known models of bilingual word recognition, viz., the Revised Hierarchical Model (RHM), the Bilingual Interactive Activation Model (BIA) and its precursor the Interactive Activation Model (IAM), the Bilingual Interactive Activation Model Plus (BIA+), and Multilink.

2.2.1. The Revised Hierarchical Model (RHM)

Kroll and Stewart (1994) – later revisited in Kroll and Tockowicz (2001) – proposed a consolidation of two hierarchical models of language interconnection conceived originally by Potter, So, von Eckhardt, and Feldman (1984) to test whether L2 proficiency could account for translation performance differences in picture naming and translation-into-L2 tasks. The first model, viz., Word Association Model, purported that words in the bilingual mental lexicon are stored in two separate, yet interconnected, lexical memory systems (one for each language), and that concepts are stored in an abstract semantic system. As Figure 1 below shows, in the Word Association Model L2 words are associated with L1 words (and vice-versa). However, only through L1 mediation can L2 words gain access to the abstract semantic system (concepts).

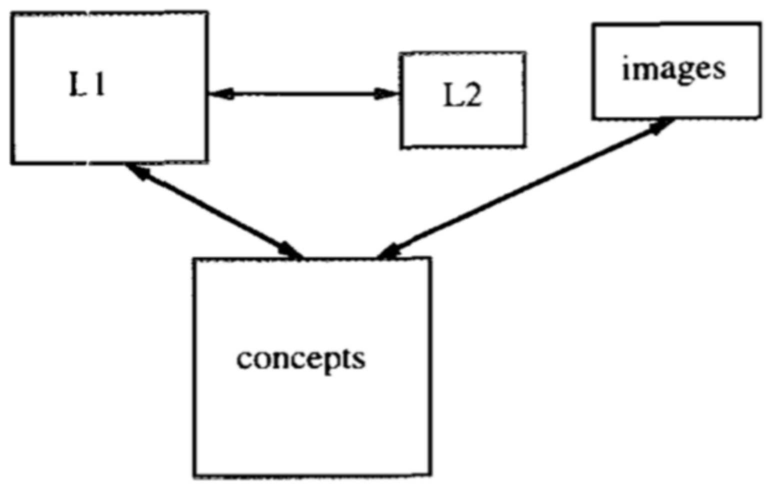


Figure 1. (1984)'s Word Association Model as depicted in Kroll and Stewart (1994, p. 150). The arrows represent lexical links.

The second model, viz., Concept Mediation Model, also proposed that words in the bilingual mental lexicon are stored in two separate lexical memory systems (one for each language) and that concepts are stored in an abstract semantic system. However, as Figure 2 shows, unlike in the Word Association Model, in the Concept Mediation Model the two lexical stores are depicted as completely independent from each other. Further, L1 and L2 words have direct access to concepts in the common abstract semantic system (concepts). However, in order for words in either lexical store to gain access to words in the other lexical store, this access would need to be mediated via concepts, hence the name Concept Mediation Model.

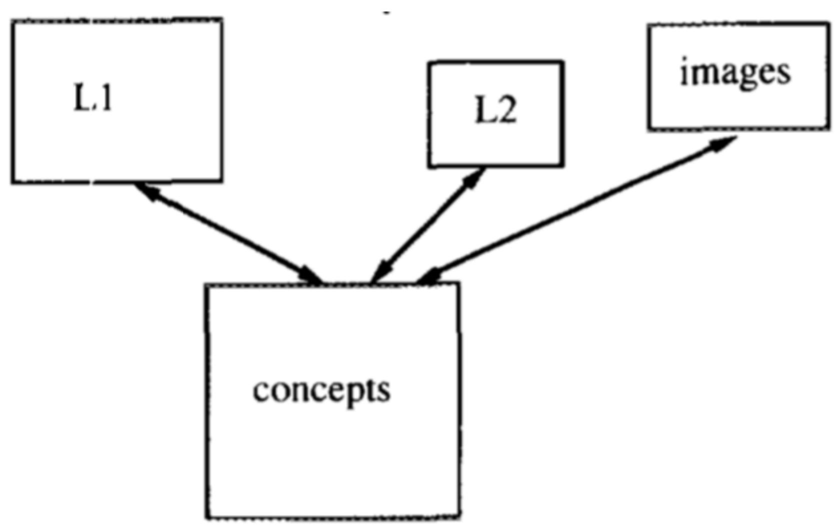


Figure 2. Potter et al. (1984)'s Concept Mediation Model as depicted in Kroll and Stewart (1994, p. 150). The arrows represent lexical links.

In a series of studies that put both models to the test, Porter et al. only found support for the Concept Mediation Model. They concluded that the semantic store must always mediate the connections between both lexical stores irrespective of bilinguals' L2 proficiency, which was challenged subsequently in Kroll and Curley (1988) as well as Chen and Leung (1989).

Similar to both the Concept Mediation Model and the Word Association Model, the Revised Hierarchical Model (RHM) also envisaged the bilingual mental lexicon following a hierarchical organization. At the lexical level, words from bilinguals' two languages are stored in their own respective lexical memory system, and at the semantic level, concepts are stored in an abstract memory system. However, unlike its two predecessor models, the RHM posits that as bilinguals become more proficient in their L2, or as bilinguals' L2 dominance in respect to L1 increases, L2 words can gain direct access to the abstract conceptual store instead of via L1 mediation.

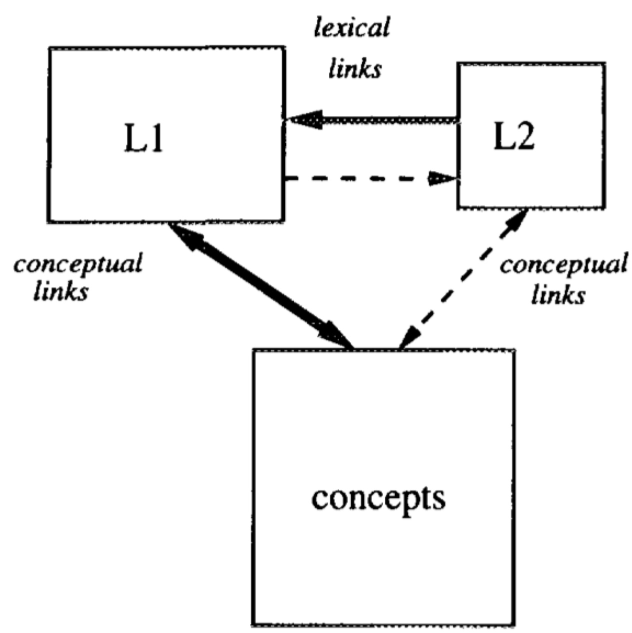


Figure 3. The Revised Hierarchical Model as depicted in Kroll and Stewart (1994, p. 158).

The RHM's organization can elegantly accommodate translation performance differences that can arise as a result of L2 proficiency (illustrated by the dashed arrows in Figure 3 above). Specifically, for unbalanced bilinguals, those who are less proficient in their L2, translating from L1 to L2 occurs through concept mediation, i.e., L1 words gain access to L2 words via the semantic store (concepts), thus, translating in this direction is slower. This conceptual mediation mirrors unbalanced bilinguals' reliance on their L1 in order to access the conceptual representations in the semantic store. Translating from L2 to L1, however, is not mediated via the semantic store, i.e., L2 words have direct access to L1 words, thus, translating in this direction is faster (illustrated by the solid arrows in Figure 3 above). This is so because the connections the L2 lexical store develops to the L1 lexical store are much stronger than the other way around, which occur through conceptual mediation.

2.2.2. The Bilingual Interactive Activation Model (BIA)

It is impossible to examine the BIA (Dijkstra & van Heuven, 1998) without first introducing the Interactive Activation Model (IAM), its precursor. The IAM is a connectionist model devised by McClelland and Rumelhart (1981; 1982) that postulates that bottom-up and top-down processes occur in parallel during the course of visual word recognition (or perception). Structurally, the IAM is a hierarchical model. It specifies various levels of representation (abstraction), viz., visual-acoustic feature level, a letter-phoneme level, a word level, and higher levels of processing that exert top-down input to the word level (see Figure 3). Low level visual and acoustic feature-based representations spread their activation to letter and phoneme-based representations at one level higher, which, in turn, spread their activation to word-based representations one level higher. As its name suggests, the

IAM posits that visual word recognition is a highly interactive process; thus, each adjacent level of representation can interact with each other, either exciting or inhibiting activation spread to other adjacent levels. However, activation spread within the word level is only inhibitory, since one word is posited to receive excitatory activation from lower levels of representation.

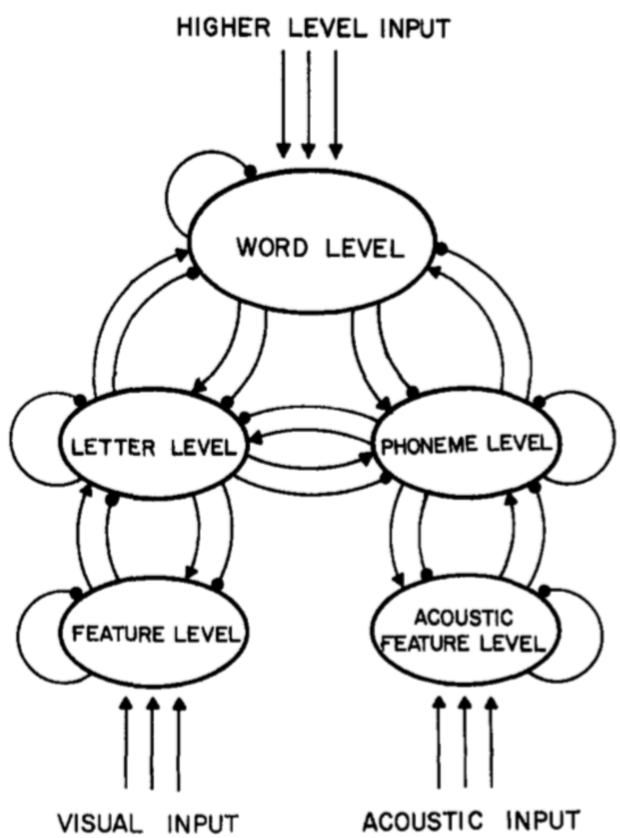


Figure 4. The Interactive Activation Model (IAM) as depicted in McClelland and Rumelhart (1981, p. 378).

The BIA is in its essence an offshoot of the IAM. Thus, its architecture is largely the same, with the exception of the added language level containing two language nodes to represent bilinguals' two languages, since the BIA is a model of bilingual visual word recognition. Activation at each level of representation in the BIA can further excite (feed-forward) or inhibit (feedback inhibition) activation at each of the adjacent levels (see Figure 5). In this way, when a string of letters is presented to the model, low-level feature-representations are activated, which subsequently activate letter-based representations with features present in the original input string and also inhibit other letter and phoneme-based representations lacking those features. The letter-based representations in turn activate words in both languages that contain the letters in the same position of the input string and also inhibit words that do not. Thus, regarding the language non-selective hypothesis, activation in the BIA is imagined to be non-language specific. Language nodes then exert top-down inhibitory feedback on the activated lexical items from both languages, inhibiting the ones from the other language to ensure successful identification of the input string.

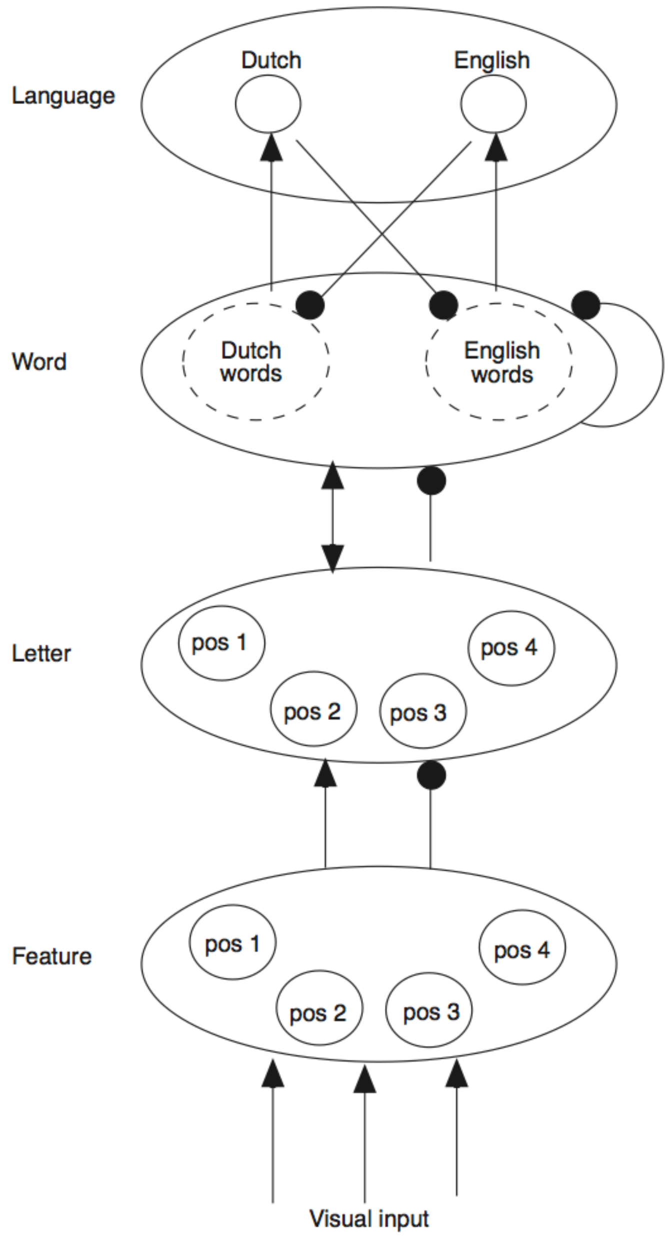


Figure 5. The Bilingual Interactive Activation Model (BIA) depicted in in Dijkstra & van Heuven (2002, p. 177). Excitatory connections are represented by arrowheads. Inhibitory connections are represented by filled circles.

2.2.3. The Bilingual Interactive Activation Model Plus (BIA+)

In order to address the lack of specification for phonological and semantic representations in the BIA, the Bilingual Interactive Activation Model Plus (BIA+) was conceived (Dijkstra & van Heuven, 2002). The BIA+, similarly to the BIA, is a computational model. It envisages bilingual word recognition being comprised of two primary systems, a Word Identification System and a Task/Decision System. Orthographic, phonological and semantic representations from bilinguals' both languages are integrated in the Word Identification System, much like in a network of interconnected nodes where sublexical and lexical linguistic representations are stored.

According to the BIA+ the bilingual mental lexicon is integrated and lexical access is language non-selective. As depicted in Figure 6, when a written word is presented to the BIA+, a number of word competitors from both languages are activated simultaneously based on their similarity to the input word. Additionally, the BIA+ postulates that bilingual visual word recognition is affected by crosslinguistic orthographic overlap as well as crosslinguistic phonological and semantic overlap. In this fashion, when a word is presented to the BIA+, its lexical and sublexical orthographic representations activate similar lexical and sublexical orthographic and phonological representations, which then activate semantic representations. Thus, the greater the degree of crosslinguistic orthographic, phonological and semantic overlap between a word and stored mental representations (in both languages), the greater the number of similar word competitors from both languages that are activated simultaneously.

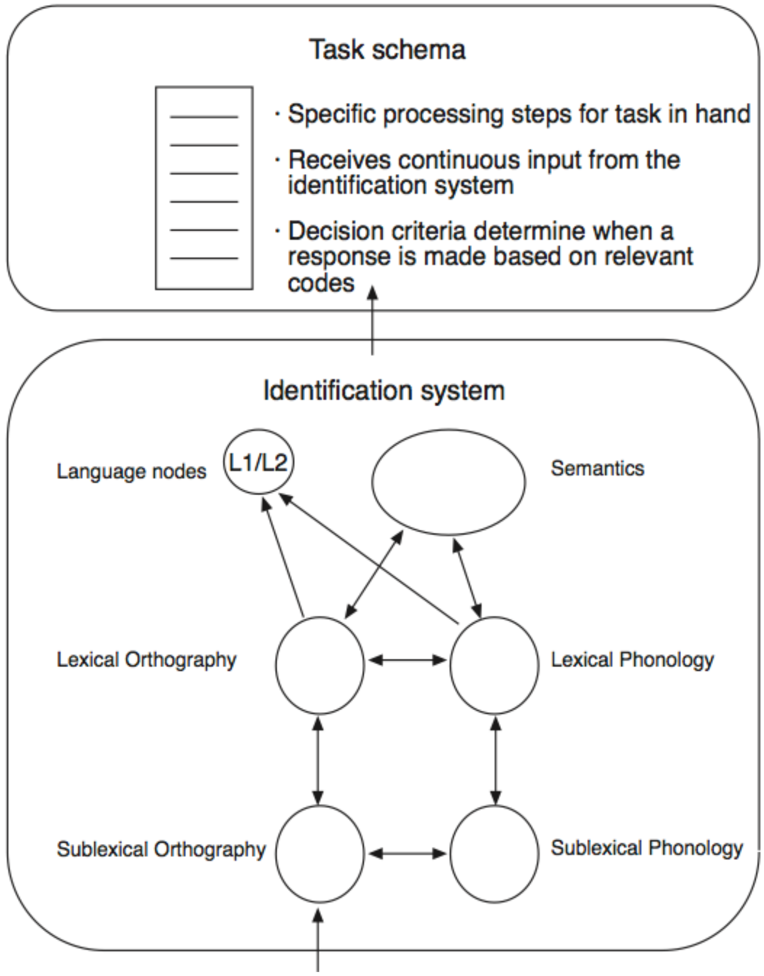


Figure 6. The Bilingual Interactive Activation Model Plus as depicted in Dijkstra & van Heuven (2002, p. 182). Activation flow is represented by arrows. Inhibitory connections are not depicted in the illustration.

The BIA+ also posits that other factors, such as word frequency, recency of use (how recently a word was last used by the bilingual) and L2 proficiency modulate activation spread crosslinguistically, because they can raise or lower the resting-level activation of stored representations. In this fashion, compared to words that are not used often, frequently used words are thought to have a higher resting-level of activation. Consequently, not only less input activation is required to activate stored representations linked to a frequently used word but also activation is believed to spread much more quickly between frequently used words. Likewise, a delay in activation spread is posited to occur among L2 representations compared to L1 representations, which is more pronounced for unbalanced bilinguals. Because this group of speakers is not equally proficient and fluent in their L2, their L2 representations are posited to have a lower resting-level of activation, thus activation spreads to linked L1 mental representations much more quickly than it does to L2 representations.

2.2.4. Multilink

Multilink (Dijkstra et al., 2018) is a state-of-the-art revision to earlier models of the bilingual mental lexicon and, as such, incorporates several features of its predecessors, in particular the RHM, the BIA and the BIA+. Multilink is a computational cognitive model based on the localist-connectionist paradigm and written in an object-oriented approach in Java. It allows precise simulation of cognitive processes using a series of computations (algorithm) that enable language researchers to generate new hypotheses for empirical testing and better understand how the mechanisms underlying bilingual (and monolingual) word recognition and production interact and play out both qualitatively and quantitatively with an unparalleled level of precision never thought possible before. Moreover, Multilink allows language researchers to advance more robust theories regarding the organization of the bilingual mental lexicon, particularly with respect to how *cognates* might be stored. Its task/decision system can replicate several experimental tasks, such as word processing during lexical decision, orthographic and semantic priming, word naming, and word translation (in both forward and backward directions), which are often utilized in psycholinguistic studies. Furthermore, Multilink's lexicon parameter can be fine-tuned to account for L2-proficiency levels, frequency of use, word length and crosslinguistic overlap. In fact, Multilink is such a new model that its lexicon at the moment only comprises English-Dutch words. Multilink's lexicon, however, is integrated across both languages. Over time, as the model becomes more mainstream with language researchers, its architects hope Multilink's lexicon will expand to include several more language pairs.

One of Multilink's core assumptions is that the mental lexicon is integrated across both languages of a bilingual. This is in line with the findings of many psycholinguistic studies suggesting there is little scientific evidence to conceive the existence of two separate lexicons, one for each language (Marc Brysbaert & Duyck, 2010; Van Heuven, Dijkstra, & Grainger, 1998). Multilink depicts the mental lexicon as layered lexical network interconnected by links that vary in strength and in which activation flow is largely bidirectional (see Figure 7 below). Nevertheless, despite having a layered structure, unlike the BIA+, a sublexical level is not specified in Multilink, thus, activation flow occurs lexically. Multilink posits each stored word as having a resting level activation (RLA), which is a function of the frequency of use of the word itself (measured in occurrences per million) and the highest and lowest frequencies of all other words stored in the lexicon. In this framework, when a written word, irrespective of its length, is presented to Multilink as input, it simultaneously activates a number of stored lexical-orthographic representations (word candidates) from both languages proportionate to the strength of the orthographic similarity between the input word and the stored representations as well as their RLA. Multilink utilizes the Levenshtein distance (Levenshtein, 1966) normalized for word length as a metric of crosslinguistic orthographic similarity. Word recognition in Multilink takes place in a number of time steps (or cycles). Thus, the degree of semantic and phonological crosslinguistic similarity between the input word and each word candidate can further strengthen or attenuate activation flow among word candidates in subsequent time cycles. However, it is important to point out that in Multilink crosslinguistic phonological similarity can only strengthen activation flow through semantics. The task/decision system lastly checks for language membership of the input word and word candidates as well as their degree of orthographic, phonological and semantic activation required for output in the experimental task at hand.

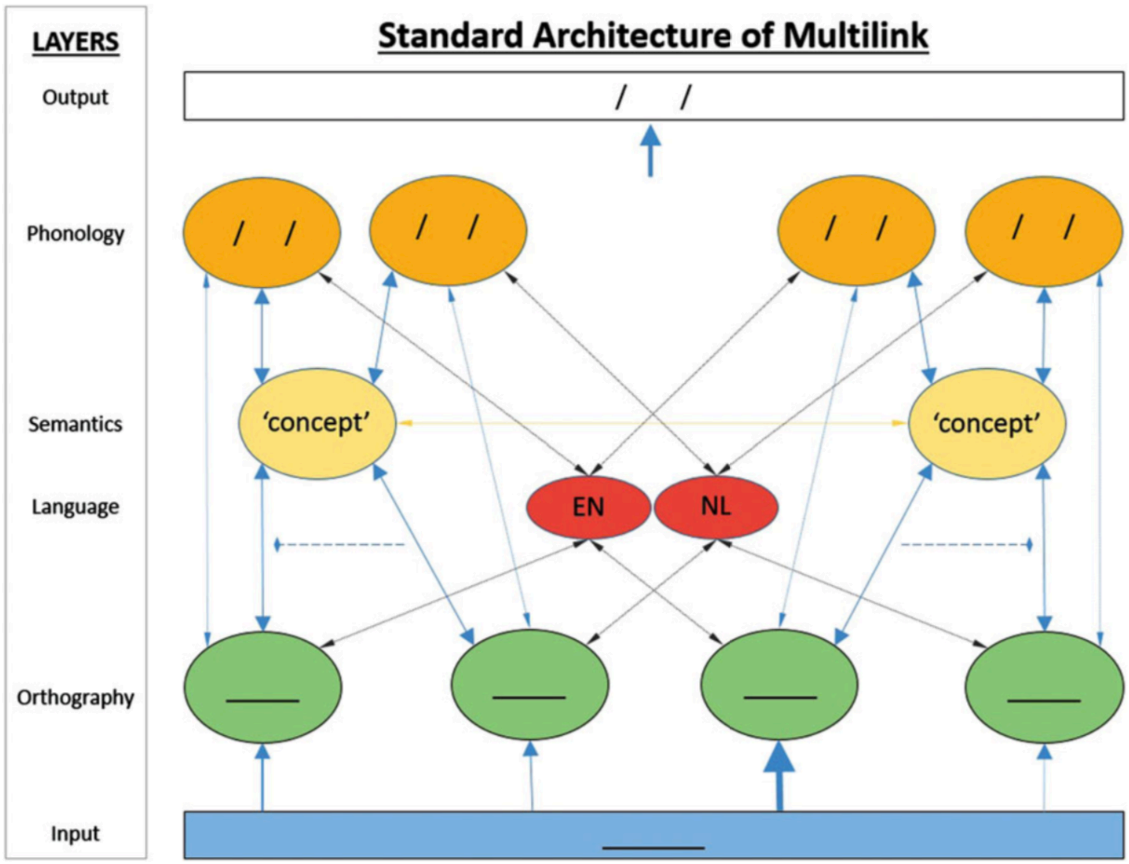


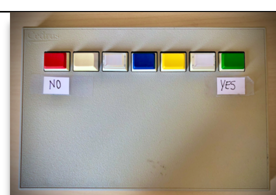
Figure 7. The network architecture of Multilink as depicted in Dijkstra et al. (2018, p. 662). EN = English, NL = Dutch. Arrows represent activation flow through the various network layers of lexical representation, from the input word (at the bottom) to output (at the top).

2.3. Testing the Organization of Bilingual Mental Lexicon: The Lexical Decision Task and The Masked-Priming Paradigm

The masked priming paradigm is one of the most well-known techniques employed by language researchers to test their hypotheses regarding the organization of the mental lexicon. It consists of sequentially presenting participants in the center of a computer screen a *forward mask*, a *prime*, sometimes a *backward mask*, and a *target*.

This paradigm allows an experimenter to manipulate the properties of the prime in order to determine its effect on participants' response latencies (or their response accuracy) to the target, thus illuminating the processing of linguistic code (Tzur & Frost, 2007). Priming occurs when the prime facilitates participants' response to the target in comparison to some neutral/control baseline (Forster, Mohan, & Hector, 2003, Chapter 1). The most common explanation for the facilitation effect induced by priming is that it is an unconscious (non-strategic) automatic process due to spreading-activation. Specifically, the prime activates a linguistic representation, and this activation spreads associatively to other linguistic representations, facilitating their identification (Hutchison, 2007). In order to measure the effect of priming, the masked-priming paradigm is combined with a lexical decision task (LDT) where participants are instructed to respond to a word (target), usually by pressing a key on a button box or keyboard, to indicate whether or not the target is a real word in a particular language, e.g. *is this an English word?* (see Table 1 below outlining the procedure).

Table 1. *LDT Trial Utilizing the Masked-Priming Paradigm*

Presentation Sequence	Computer Screen Display	Duration
1 forward mask hash marks covering the same visuospatial field of the prime	######	500ms
2 prime in lowercase letters	fruta	30-70ms SOA
backward mask (optional)	######	500ms
3 target in uppercase letters	FRUIT	500ms
4 participant's response usually pressing a key on a button box (yes/no)		3,000ms before the next stimulus is presented

As Table 1 outlines, a number of procedural parameters needs to be carefully considered prior the design stage and during the administration of an LDT. A particular important one is the *stimulus onset asynchrony* (SOA), which is the amount of time that will elapse between the presentation of the prime and subsequent target. A shorter SOA invokes early word recognition processes and ensures that participants are completely unaware of the prime, thus mitigating the likelihood that any cognitive processes (linguistically related and potentially non-linguistically as well) are triggered (Forster & Davis, 1984; Forster et al., 2003; Grainger & Holcomb, 2009; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Sabourin, Brien, & Burkholder, 2014). Long SOAs make the prime completely visible to participants, i.e., they are able to completely read the prime through, and, since they are conscious of it, postperceptual cognitive processes, not all connected to the language system, are automatically induced. (Forster et al., 2003, Chapter 1). Specifically, long SOAs allow participants to consciously (and strategically) compute probable targets (conscious expectancy generation). As a result, long SOAs yield rather limiting priming effects, because other cognitive operations interfere with the time course of the processing of a linguistic code (Colombo, 1986; Martin & Jensen, 1988; Tzur & Frost, 2007).

As can be seen in Table 1, a trial in an LDT is comprised of four major events that occur in the following sequence: (1) a forward mask consisting of a sequence of hash marks/tags (#####) covering the same visuospatial field as the prime is presented for 500ms before the prime; (2) the prime in lowercase letters is presented briefly (SOA 30-70 ms); (3) the target, usually in uppercase letters, is presented; and (4) participants have a preset amount of time to record their response (usually by the press of a button on a button box) before the next trial is displayed. It is important to highlight the case change between the prime and target, for it ensures that the two stimuli are visually distinct. Additionally, in order to avoid participants perceiving the target as a continuation of the prime due to their form overlap, some studies have opted to add a backward mask between the presentation of the prime and the subsequent presentation of the target (Forster et al., 2003, Chapter 1).

Distinct types of priming effects are reported in the literature when the masked-priming technique is used in experimental tasks and have helped researchers to postulate and advance models of visual word recognition and lexical organization. “These effects are assessed relative to a baseline condition, in which the prime differs from the target in all baseline conditions” (Forster et al., 2003, para. 6). Among the effects observed, the *translation priming effect* is of particular interest because the prime in one language is posited to facilitate the processing of a translation equivalent target in the other language, e.g., the Portuguese prime *cão* would facilitate the processing of its English translation equivalent *dog*. This effect is attributed to the mental representations of both words being connected in memory either at a conceptual or a lexical level. In addition, *translation priming asymmetry* has been reported across several different language pairs, and is significantly greater in the forward priming direction, viz., from L1-to-L2, than in the backward direction, viz., L2-to-L1 (Duñabeitia, Dimitropoulou, Uribe-Etxebarria, Laka, & Carreiras, 2010; Gollan, Forster, & Frost, 1997; Grainger & Frenck-

Mestre, 2002; Jiang, 1999). In the forward-priming direction, L1 translation primes consistently yield faster reaction times to L2 targets, with an average magnitude effect of 39ms. In contrast, in the backward-priming direction, the results are not as consistent, with an average magnitude of the effect remarkably less robust at around 6ms (Duñabeitia, Dimitropoulou, et al., 2010). This priming asymmetry is thought to be the result of uneven proficiency between the L1 and the L2, i.e., bilinguals who are not equally proficient in both of their languages (Duñabeitia, Dimitropoulou, et al., 2010; Gollan et al., 1997; Grainger & Frenck-Mestre, 2002; Jiang, 1999). Nevertheless, more recently Wen and van Heuven (2017) conducted a meta-analysis of 24 studies that have utilized the masked-priming technique to quantitively assess the effect sizes of translation priming effects in both directions. The results revealed significant priming effects in both priming directions, with the effect being significantly larger in the L1-L2 than the L2-L1 direction.

The *semantic priming effect* is attributed to the automatic activation of a shared semantic store, i.e., the prime in L1 and the target in L2, or vice versa, share a single conceptual (semantic) representation. In this way, when an L1 prime is presented to a bilingual participant, it automatically activates a target in L2, or vice-versa (Dimitropoulou, Duñabeitia, & Carreiras, 2011a; Van Hell & Dijkstra, 2002). For example, for a Portuguese-English bilingual a Portuguese prime like *café* 'coffee' is expected to activate English targets like *milk* or *sugar*, since coffee is usually served with milk and sugar in many Portuguese speaking countries.

The *orthographic priming effect* (also described as *form priming*) is a type of priming effect reported by studies that have utilized the masked-priming paradigm. The effect is reported when the processing of a target, e.g., faster reaction times, is facilitated by the presentation of a preceding orthographically similar prime. In this condition, when the orthographic overlap between the prime and target is high, i.e., when they share many letters, there is minimal competition between them, thus perception of the target is facilitated.

The *phonological priming effect* is another priming effect that has helped enhance several models of bilingual visual word recognition and is often reported in the literature. This effect is of particular interest because it is often present for bilingual participants but absent for monolingual participants. A few studies that have utilized the masked-priming technique to investigate whether it is possible for a homophonic prime in one language to facilitate the recognition of a target in the other language have suggested that during the initial stages of visual word recognition, in addition to its orthographic representation in both languages, a word's phonological representation (in both languages as well) is also automatically activated (Marc Brysbaert, Van Dyck, & Van De Poel, 1999; Dimitropoulou, Duñabeitia, & Carreiras, 2011b; Perfetti & Bell, 1991; Pexman, Lupker, & Jared, 2001; Van Wijnendaele & Brysbaert, 2002). "Visual word recognition is phonologically mediated to a large extent, not only in L1 but also in L2" (Marc Brysbaert, 2003, p. 186). Therefore, recognition of an L2 target, for example, can also be facilitated if a homophonic[2] prime in L1 is presented to bilingual participants.

While employing the masked priming technique has been proven useful at establishing a richer understanding of bilingual lexical organization, estimating the role of crosslinguistic phonological similarity in isolation requires a much more refined approach from language researchers. It is methodologically challenging to disentangle orthographic similarity from phonological similarity due to the fact that in most languages letters (orthography) map to sounds (phonemes). More often than not, words that are orthographically similar crosslinguistically (have large crosslinguistic orthographic code overlap/similar spelling) are also judged to be phonologically similar crosslinguistically (large phonological code overlap/similar pronunciation). Simply put, when a word in one language sounds similar to a word in another language, their spelling is also similar. Nevertheless, newer models of bilingual visual word recognition, viz., BIA+ and Multilink, have incorporated crosslinguistic phonological similarity as crucial component separate from crosslinguistic orthographic similarity.

---

[2] The word "homophonic" is used here with the meaning of sounding similar or having a similar pronunciation.

2.4. Cognates and false friends

Many well-known models of the bilingual mental lexicon theorize that bilinguals' two languages share a single semantic store and that words are comprised of abstract linguistic representations. If the premise of a shared semantic store is accurate, due to spreading activation, during visual word recognition, presenting an L1 prime to a bilingual should automatically activate an L2-semantically-related target (or vice-versa). In addition, crosslinguistic form overlap (phonological and orthographic similarity) should also play a role. With this in mind, the masked-priming paradigm has been extensively used to elucidate the organization of the bilingual mental lexicon. To that end, two classes of words, viz., cognates and false friends, have been under close inspection by language researchers that, for the past few decades, have persistently tried to further their understanding of how L1 and L2 words are stored and organized in the bilingual mental lexicon. *Cognates* are words whose meaning is largely the same crosslinguistically, i.e., have near complete semantic overlap. Likewise, their orthographic overlap (form similarity) is also high, sometimes complete. However, their phonological overlap (pronunciation similarity), which sometimes can be high, is subject to more variation crosslinguistically (da Luz Fontes & Schwartz, 2010; Dijkstra, Grainger, & Van Heuven, 1999; Dijkstra & van Heuven, 2002; Schepens, Dijkstra, Grootjen, & van Heuven, 2013). For example, the English word kit [kɪt] and the Portuguese word *kit* [ˈki.tʃi] have complete crosslinguistic semantic and orthographic overlap.

Specifically, their meaning is largely the same and their orthographic form is identical across both languages. However, there is variation in how speakers of Portuguese and English pronounce both words. Like cognates, *false friends* (often characterized as *interlingual homographs*[3]) have a high degree of orthographic overlap and varying degrees of phonological overlap crosslinguistically. However, unlike cognates, false friends denote different concepts across languages, i.e., their semantic overlap is rather low or null (da Luz Fontes & Schwartz, 2010; Dijkstra et al., 1999; Dijkstra & van Heuven, 2002; Schepens et al., 2013). For example, while the English word *rim* [ɹĩm] and the Portuguese word *rim* [χĩŋ] 'kidney' have complete orthographic overlap and partial phonological overlap, their meanings are completely different across both languages.

In the research literature cognates appear to have a clear processing advantage. They generally yield significantly more robust and more consistent priming effects (de Groot & Nas, 1991; Gollan et al., 1997; Kim & Davis, 2003, Sánchez-Casas, Davis, & García-Albea, 1992). Cognates and false friends, either as primes or targets in LDTs, have often been compared against controls and pseudowords (Dimitropoulou et al., 2011b; Ferrand & Grainger, 1993; Grainger, Kiyonaga, & Holcomb, 2006; Grainger, Spinelli, Farioli, Diependaele, & Ferrand,

---

[3] Interlingual homographs are sometimes referred to in the literature as *interlingual pseudohomophones* (Dijkstra et al., 1999).

2003; Perfetti & Tan, 1998; Pollatsek, Perea, & Carreiras, 2005; Ziegler, Ferrand, Jacobs, Rey, & Grainger, 2000). *Pseudowords* are words that have been manipulated either via computer software or manually. They resemble real words in a specific language in terms of both their form and pronunciation.

When orthographic overlap is the primary focus of research, facilitatory effects, viz., faster reaction times, have been reported for cognates with *partial* as well as *100%* orthographic overlap, e.g., the Dutch-English cognate pair *tomaat-tomato*, in comparison to control words (Cristoffanini, Kirsner, & Milech, 1986; Dijkstra et al., 1999; Lemhöfer & Dijkstra, 2004; Van Hell & Dijkstra, 2002). When cross-linguistic phonological overlap is investigated, such as when cognates are pronounced similarly across two languages, evidence from LDTs shows faster reaction times as well, suggesting that phonological representations might be activated during visual word recognition, and, consequently, might play a role in the organization of the bilingual mental lexicon. This facilitatory effect is observed even with languages pairs that do not share the same orthographic script, such as Greek and Spanish, Japanese and English, as well as Greek and French (Dimitropoulou et al., 2011a; Nakayama, Sears, Hino, & Lupker, 2012; Voga & Grainger, 2007).

Costa, Caramazza and Sebastián-Gallés (2000) theorize that complete semantic overlap between a prime in one language and a target in the other could potentially serve as proof that activation spreads backward from semantic to orthographic (and potentially to phonological) representations, accounting for some of the facilitatory effects encountered in the literature. Nevertheless, Voga and Grainger (2007) present a different theory. Specifically, they argue that the cognate advantage over noncognates may simply be the result of strong morphological, orthographic and phonological relationships that exist between primes and targets across some language pairs. It is possible that words that share a common etymological root also share a common morphemic representation in memory, thus, are organized differently in the bilingual mental lexicon.

In sum, research on cognates has thus far provided, to a certain extent, valuable evidence for language non-selectivity, i.e., bilinguals' two languages appear to be simultaneously activated during visual word recognition. This parallel activation points to a continuous state of co-activity in the bilingual mental lexicon, in which not only semantic representations but also orthographic and phonological representations are also activated as a result of spreading activation when a word from one of bilinguals' two languages is presented (de Groot, 1992; J. F. Kroll & Stewart, 1994; Perea, Duñabeitia, & Carreiras, 2008; Soares & Grosjean, 1984; van Hell & de Groot, 1998). Nevertheless, teasing apart the individual contributions of crosslinguistic semantic, orthographic and phonological similarity has remained a challenge due to the vastly different methodologies employed by language researchers.

2.5. The Current Study

In light of the framework provided primarily by the BIA and BIA+, and to some extent Multilink, the primary goal of this study is to quantify the individual effects of crosslinguistic orthographic, phonological and semantic similarity on the organization of the bilingual mental lexicon. In addition, this study seeks to understand whether language proficiency plays a significant role in this organization. To accomplish that, Portuguese-English cognates and false friends were selected as experimental stimuli for a cross-language LDT with masked priming, for which Portuguese-English bilinguals and English functional monolinguals were selected to participate.

To date no study has concerned primarily with Portuguese-English cognates and false friends in a cross-language LDT, despite the great deal of crosslinguistic similarity between the two languages. As Brazil, the most populous Portuguese speaking country, has become one of the world's most powerful economies, consistently ranking among the top 10, and experienced dramatic economic growth, the country's participation in the world's economic stage has also become more prominent ("The World's Top 10 Largest Economies," n.d.). As a result, the number of native speakers of Portuguese receiving training to become Portuguese-English bilinguals has also increased. Thus, a more scientific understanding of how words from two languages are stored and interact is needed and it will certainly assist in the development of better teaching materials and linguistic outcomes.

Despite the fact that the likelihood that two words across two different languages are cognates is highly correlated with how similarly they are pronounced (Kondrak & Sherif, 2006), crosslinguistic phonological similarity has not received much attention from language researchers; thus, its individual effect on bilingual mental organization has not been clearly established or understood (Dimitropoulou et al., 2011a; Lemhöfer & Dijkstra, 2004; Nakayama et al., 2012; Van Hell & Dijkstra, 2002; Voga & Grainger, 2007).

Because there currently is no standard crosslinguistic phonological similarity metric readily available for language research, researchers have opted to reduce crosslinguistic phonological similarity to segmental similarity and explain it in terms of distinctive features and other formal primitives of phonological theory (Chomsky & Halle, 1992; G. Clements & Hume, 1995; G. N. Clements, 1985; inter alia). In fact, influential studies on the bilingual mental lexicon have, for the most part, either relied on this traditional framework for assessing crosslinguistic phonological similarity or on native and non-native speakers' personal intuitions for selecting cognates as stimuli for their experiments (Comesaña et al., 2012; Friesen & Jared, 2012; Pallier, Colomé, & Sebastián-Gallés, 2001; Pexman, Lupker, & Jared, 2001; Schwartz & Kroll, 2006; Schwartz, Kroll, & Diaz, 2007; inter alia).

Unfortunately, employing subjective measures of crosslinguistic phonological similarity pose numerous challenges and limit the results of these studies, in particular, because the mechanisms involved in speech perception and production are far more multidimensional than what is postulated by phonological theory and personal judgments are not only often biased but also based on the personal experiences of the speakers of the languages being studied (Bradlow, Clopper, Smiljanic, & Walter, 2010). These pitfalls, however, can be successfully overcome if an objective metric is used instead. To that end, dynamic time warping (DTW), an algorithm that was originally conceived for automatic speech recognition to measure acoustic similarity between two waveforms, was leveraged to create an objective crosslinguistic phonological similarity metric. This approach is largely an offshoot of the technique developed by Mielke (2012).

To better understand the individual effect of crosslinguistic orthographic similarity as well as its interplay with crosslinguistic phonological similarity on bilingual lexical organization, it is essential that an objective metric of crosslinguistic orthographic similarity also be used. To that end, in light of the work of Schepens, Dijkstra and Grootjen (2013), the Normalized Levenshtein Distance (NLD) was chosen.

2.5.1. Experimental Predictions and Hypotheses

Taking into account spreading activation as well as the framework of two well-known models of the bilingual mental lexicon discussed in 2.2., specifically the BIA and BIA+, this study hypothesizes that activation in the bilingual mental lexicon spreads in a feed-forward fashion from the semantic store, which is the site where bilinguals' two languages are posited to be integrated, to the orthographic representations from both bilinguals' two languages during visual word recognition. Phonological representations are activated via orthographic representations, not through the semantic store. In this framework, cognates will have a clear processing advantage over false friends. Specifically, Portuguese-English bilinguals will recognize an English target significantly faster when it is matched to a Portuguese cognate prime compared to when it is matched to a Portuguese false friend prime. The shared semantics coupled with orthographic and phonological code overlap will accelerate the recognition of English cognate targets. When an English target is matched to a Portuguese false friend prime, the orthographic overlap and phonological overlap will still facilitate the processing of the English target, but the advantage will be less pronounced than for Portuguese cognate primes due to no semantic overlap. Bilingual participants should recognize English targets matched to a semantically unrelated Portuguese prime with low or null orthographic and phonological overlap with the English target (control pairs) slower because the orthographic and phonological code overlap is minimum or absent. Finally, bilingual participants will recognize pseudoword targets matched to a Portuguese prime (distractor pairs) the slowest due to the absence of facilitation provided by semantic, orthographic and phonological overlap. These predictions are detailed in Table 2 below.

Table 2. Experimental Predictions (Hypotheses)

Prime Category	Crosslinguistic Overlap with English Target			Bilinguals' Recognition of English Targets (Reaction Time)
	Semantic	Orthographic	Phonological	
Portuguese *cognate*	complete	partial or complete	partial or complete	fastest
Portuguese *false friend*	absent	partial or complete	partial or complete	second fastest
Portuguese word matched with unrelated English target (*control pairs*)	absent	minimal or none	minimal or none	slower
Portuguese word matched with pseudoword (*distractor pairs*)	absent	none	none	slowest

In the extant literature, ANOVAs and t-tests have been the primary statistical analysis (means-based parametric) tools used to analyze bilingual data collected in LDTs, and the reliance on these tools has continued to this day. Because one of the primary goals of this study is to quantify the individual effects of crosslinguistic similarity and because many of the variables of interest are continuous predictors, or a combination of categorical and continuous predictors, this study will instead utilize linear mixed-effects models, a much more robust statistical analysis tool, to analyze the experimental data collected. Besides more precisely estimating the individual effects of crosslinguistic similarity (fixed factors) on both dependent variables (reaction time and accuracy), mixed-effects models will enable participants and prime-target pairs (items) to be treated as random effects, which will allow the individual effects of crosslinguistic similarity to be further assessed by each individual participant and each individual prime-target pair, which in the end will yield a more realistic overall representation of the processes involved in visual word recognition in the bilingual mental lexicon.

The following chapter provides a thorough account of the study's methodology, from the creation of the stimuli lists for the LDT using objective metrics of crosslinguistic orthographic and phonological similarity to recruiting both Portuguese-English bilinguals and functional English monolingual participants as well as assessing their level of English proficiency. It also describes the experimental procedure used in the cross-language LDT.

3. The Study: Methodology

3.1. Participants

3.1.1. Recruitment

Participants for the experiment were recruited primarily via a recruitment email,[4] which provided potential participants with a brief description of the experiment, qualification requirements, a step-by-step account of what their participation would entail, and instructed them to contact the researcher via email to express their interest in serving as volunteers. In order to recruit Portuguese-L1 English-L2 late sequential bilinguals, the researcher sent out the recruitment email (see Appendix B) to the International Office at the University of Ottawa, and requested that they forward the message to any current international graduate or undergraduate student whose native language is Portuguese. Since a large proportion of students with this language profile conducts their studies at the University in English, they were presumed to possess a level of English that allows them to function in an academic environment, which qualified them to participate in the experiment.

To serve as *functional monolingual controls*, the researcher sent out the recruitment email to the teaching staff of the English Intensive Program (EIP), a non-credit English as a second language program in the Official Languages and

[4] The text included in recruitment email was approved by the University of Ottawa Office of Ethics and Research Integrity (file number 09-12-13). The approval notice is included in Appendix A.

Bilingualism Institute at the University of Ottawa. The EIP is designed for international students and francophone Canadians who do not meet the English proficiency requirements of the University so that they might attend classes either at the graduate or undergraduate level. Teachers in the EIP are highly trained professionals who must at a minimum have a Masters in linguistics, second language education, or related disciplines in order to be a member of the teaching staff, and many of them are native speakers of English.

It is important to note that finding potential monolingual participants in the National Capital Region in Canada is a rather dauting endeavor, as residents of the area are exposed to both official languages, English and French, on a daily basis. Street signs, advertisements, pamphlets, bus schedules, routes, etc. are omnipresent in both languages. To further complicate matters, Canada has a rather ethnically diverse and multicultural population, so many native speakers of English are also heritage speakers of another language, i.e., they can converse with their parents and family members in another language with varying degrees of fluency. Thus, it is virtually impossible to find bona-fide native speakers of English who do not know, speak, and are exposed to another language to serve as monolingual controls. To minimize the effects of exposure and having some knowledge of a second language, potential English monolingual participants were only allowed in the experiment if they disclosed during a prescreening interview that they could not comfortably function in an environment that would require them to use their second language in the same manner as they would English.

Many participants during the prescreening interview indicated that they knew other potential participants that met the language requirements of the experiment and who would probably be interested in serving as volunteers. Therefore, a few participants in both experimental groups actually became aware of the experiment via word-of-mouth.

3.1.2. Composition

In total 56 participants, specifically, 29 L1-Portuguese L2-English late sequential bilinguals as well as 27 L1-English functional monolinguals, participated in the experiment. All of them were 18 years of age or older at the time of their participation. All bilingual participants were born and grew up in country which Portuguese is the official language and indicated they had learned English as a second language in private language schools in their home country during or after reaching puberty. A considerable number of bilingual participants from Brazil were enrolled at the University of Ottawa under the Science Without Boarders program, a large-scale nationwide scholarship program primarily funded by the Brazilian federal government. The program seeks to strengthen and expand the initiatives of science and technology, innovation and competitiveness through international mobility of undergraduate and graduate students and researchers. Many of the students enrolled under this program at the University of Ottawa obtained scores in the Test of English as a Foreign Language (TOEFL) or the International English Language Testing System (IELTS) that satisfied the English proficiency requirements of the University. Students who did not take or obtained a low score either of these tests had to complete a few sessions in the EIP in order to satisfy the English proficiency requirements of the University before they could start taking classes at either the undergraduate or graduate level.

Most participants from Brazil were from the State of São Paulo. All monolingual participants were born and grew up in a country in which English is one of the official languages.

3.1.3. Language Background Profile and Language Proficiency Assessment

In order to establish participants' language background profile, all participants completed an abridged version the Language Background Questionnaire (see Appendix D), which was originally developed for the psycholinguistic experiments conducted in the Brain and Language Laboratory in the Department of Linguistics at the University of Ottawa and was previously used in Alves-Soares (2013). The questionnaire was converted to *Google Forms* so that participants could fill it out online on a computer and to make it convenient for them to complete it in a timelier fashion. To assess participants' overall level of proficiency in English, consistent with the methodology employed in Alves-Soares (2013), all participants were also required to take an integrative language test, specifically Brown's (1980) cloze test (see Appendix E). This kind of test was selected for the purposes of this experiment because it is easily administered and provides an overall picture of communicative competency and overall linguistic ability. In addition, cloze tests require that test takers make highly complex series of grammatical and lexical decisions, specifically having to deduce the deleted words based upon all available contextual clues, which are posited to reflect test takers' acquired language skills (Litz & Smith, 2006).

Brown's (1980) cloze test consists of a short passage of approximately 400 words in length and with 50 missing words or blanks entitled *Man and His Progress* (see Appendix E). Participants were instructed to first read the passage as a whole to get the general meaning and then fill in the blanks with the words they judged appropriate for the context. They were given 30 minutes complete the test.

Participants' answers were scored using both the semantically acceptable scoring procedure (SEMAC) and the exact replacement scoring method (ERS). In the SEMAC procedure, the test taker's answers are compared with the original deleted words. If the supplied words preserve the original intended meaning, even if they are not direct synonyms of the deleted words, they are deemed acceptable and marked as correct answers. In the ERS method only supplied words that are exactly identical to the deleted words are marked as correct answers. It is important to note that

SEMAC and ERS scores correlate very highly with one another (Litz & Smith, 2006). Table 3 below provides further details about the participant population.

Table 3. Participant Population Breakdown

Native Language	**Country of Origin**	**N**	**Age Mean (with Range)**	**Cloze Proficiency Score Mean (with Std. Dev.)**	
				ERS	**SEMAC**
Portuguese	Brazil	26	33 (23-56)	49.67 (11.91)	75.83 (16.92)
	Portugal	2	36 (33-39)	43.00 (1.00)	78.00 (6.00)
	The Azores	1	41	68.00	78.00
English	Canada	23	38.65(25-66)	62.23 (9.43)	92.46 (5.33)
	United States	4	54.75 (46-64)	70.50 (2.19)	94.50 (2.59)

3.2. Materials/Stimuli

The stimuli list for the experiment was created in six stages: (1) compiling an experimental items list comprised of Portuguese primes - English targets; (2) matching each English target with an unrelated Portuguese prime to serve as controls; (3) compiling a list of pseudowords and matching each of them with a semantic and orthographic unrelated Portuguese prime to serve as distractors; (4) compiling a list of practice pairs; (5) obtaining a frequency-per-million-words count for each Portuguese and English word; (6) obtaining an orthographic and an acoustic similarity overlap value for each prime-target pair.

3.2.1. Assembling Word Lists A and B

3.2.1.1. Experimental Items List

This study is primarily interested in two classes of word with respect to their form and meaning overlap across languages: *cognates* and *false friends*. As indicated in 2.4, cognates are words whose meaning remains conceptually stable crosslinguistically, i.e., have near complete semantic overlap. In addition, their spelling and sound similarities (orthographic and phonological overlap, respectively) are also high, sometimes complete (da Luz Fontes & Schwartz, 2010; Dijkstra et al., 1999; Dijkstra & van Heuven, 2002; Schepens et al., 2013). For example, the Portuguese word *prisão* [pri.ˈzɐ̃ʊ̃] forms a cognate pair with the English word *prison* [ˈprɪ.zə̃n], since their semantic, orthographic and phonological overlap remains stable across both languages. *False friends,* in contrast, are words that very much resemble cognates in terms of form similarity, i.e., they also possess a high degree of orthographic and phonological overlap, but they denote different concepts across languages (da Luz Fontes & Schwartz, 2010; Dijkstra et al., 1999; Dijkstra & van Heuven, 2002; Schepens et al., 2013). For example, the Portuguese noun *sapo* [ˈsa.pʊ] 'frog' shares many of its letters and phonemes with the English noun *sap* [sæp] 'fluid that circulates in the vascular system of a plant.' However, their form similarities are only superficial because the two words ultimately represent different concepts in both languages.

For the current study, an experimental items list comprised of Portuguese–English cognates and false friends was assembled using the official word list utilized in Alves-Soares (2013) as its foundation. The experimental stimuli on that list were comprised of Portuguese-English cognates and false friends. Specifically, Portuguese words served as primes; English words served as targets. As expected, the selected cognates and false friends on that earlier list had various degrees of crosslinguistic orthographic and phonological overlap.

To enhance the earlier list utilized in Alves-Soares (2013), an additional number of Portuguese-English cognates and false friends with various degrees of crosslinguistic orthographic and phonological overlap were chosen in consultation with native speakers of Portuguese who speak English as a second language (Portuguese-English sequential bilinguals). The pairs from both lists were then consolidated into a master Excel spreadsheet and classified as either cognates or false friends, which became the experimental items list for the current study. Table 4 below provides a sample of the experimental items list. The full list is provided in Appendix F.

Table 4. Portuguese primes – English targets (examples)

Classification	Portuguese primes		English targets
Cognates	alarme		alarm
	bagagem		baggage
	coma		coma
	computador		computer
	dama		dame
	face		face
	gelatina		gelatin
	irado		irate
	memória		memory
	normal		normal
	paciência		patience
	papel		paper
	robô		robot
	veia		vein
	zebra		zebra
False Friends	batom	*'lipstick'*	baton
	braço	*'arm'*	brass
	carta	*'letter'*	cart
	chato	*'boring/annoying'*	chat
	êxito	*'success/achievement'*	exit
	fábrica	*'factory'*	fabric
	gripe	*'cold/flu'*	grip
	hospício	*'psychiatric hospital'*	hospice
	limpo	*'clean'*	limp
	mal	*'bad/disorder'*	mall
	máscara	*'mask'*	mascara
	rim	*'kidney'*	rim
	sapo	*'frog'*	sap
	sorte	*'luck'*	sort
	tampão	*'lid'*	tampon

3.2.1.2. Control Items List

To serve as experimental controls, in consultation with native speakers of Portuguese who speak English as a second language (Portuguese-English sequential bilinguals), each of the Portuguese words in the experimental items list was matched with a semantically unrelated English word that did not orthographically resemble the Portuguese word, i.e., had a different spelling. For example, the Portuguese word *rim* 'kidney' in the false friend pair *rim-rim* was matched with the semantically and orthographically unrelated English word *sun,* yielding the control pair *rim-sun*. The full list of control pairs is provided in Appendix F.

3.2.1.3. Distractor Items List

Wuggy (Keuleers & Brysbaert, 2010) a multilingual pseudoword generator, was used to generate pseudowords that resembled genuine English words as closely as possible. The pseudowords were then matched with Portuguese words that were not in the experimental and control items lists. The total number of Portuguese prime – pseudoword target pairs matched the total combined number of experimental and control pairs together. The full list of distractor pairs is provided in Appendix F.

3.2.1.4. Practice Items List

To familiarize participants with the experimental task and before each trial began, practice items were used. To that end, 36 practice pairs were created. Thirtysix words of Portuguese that had not been previously used as experimental, control or distractor items served as primes. They were matched with 18 unrelated words of English that had not been previously included in the preceding lists. The remaining 18 Portuguese primes were matched with pseudowords that were generated by *Wuggy*. The full list of practice pairs is provided in Appendix F.

3.2.1.5. Official Word Lists A and B

Once the creation of the experimental, control, distractor and practice items lists was complete, two official word lists (A and B) were created for the experimental task. One half of the experimental and control pairs as well as all of distractor and practice pairs formed list A. Conversely, the second half of the experimental and control pairs as well as all of the distractor and practice pairs formed list B. It is important to note that while all distractor and practice pairs in both lists were exactly the same, the experimental and control pairs in list A were different in list B. As Table 5 shows, each official word list contained a total of 624 Portuguese prime - English target pairs, specifically, 147 experimental, 147 control, 294 distractor and 36 practice pairs.

Table 5. Word Lists Breakdown

Word List	**Prime-Target Pairs**		**N**
A	Experimental	Cognates	83
		False friends	64
	Control		147
	Distractor		294
	Practice		36
Total			624
B	Experimental	Cognates	83
		False friends	64
	Control		147
	Distractor		294
	Practice		36
Total			624

All the experimental and control pairs from both word lists as well as all the distractor and practice pairs were consolidated into a master Excel spreadsheet. This was done to further facilitate data entry, which will be explained in the following subsections. See Appendix F for the complete word lists.

3.3. Obtaining Word Frequency Count

A word frequency count was obtained for each Portuguese-English cognate and false friend in the official word lists. This was done because word frequency has major implications for this study. Specifically, the frequency in which speakers of a language are exposed to words, particularly in reading, may shape the organization of their mental lexicon. Words that are more frequently encountered are more salient to speakers of that language, which probably increases their baseline level of activation. In addition, since a vast number of words in Portuguese and English are of Latin roots, it is possible the effects of word frequency interact with the degree of orthographic (and possibly phonological) overlap.

A word frequency count is obtained by measuring how often a particular word occurs per million words in collected corpora (Schepens et al., 2013). To that end, to obtain the frequency of occurrence of each word in the preliminary word list *O corpus do Português*[5] (Davies, 2017b), a large online corpus of Portuguese words, was used to obtain word frequency count for the Portuguese words, while *Corpus of Contemporary American English (COCA*[6]*)* (Davies, 2017a), was used to obtain word

[5] *O Corpus do Português* is a one-billion-word corpus of Brazilian and European Portuguese words. It was created by Mark Davies (Brigham Young University) and funded by the US National Endowment for the Humanities (Davies, 2017b).

[6] *Corpus of Contemporary American English (COCA)* is the largest freely available and balanced corpus of American English. It contains over 560 million words from spoken, fiction, popular magazines, newspapers, and academic texts (Davies, 2017a).

frequency count for the English words. The word frequency count for each of the Portuguese-English (prime-target) word pairs was then entered into in the master spreadsheet in two separate columns, one for Portuguese and one for English. It is important to note that no word frequency count was obtained for the pseudowords (distractor items) in this experiment, since the selected pseudowords are not actually real, bona fide English words.

3.4. Obtaining Crosslinguistic Orthographic Similarity Estimates

Crosslinguistic orthographic similarity is concerned with how much spelling similarity, how many letters are shared between word pairs from different languages. Thus, orthographically similar words, also referred to as *homographs,* have a high degree of orthographic overlap (Schepens et al., 2013). The degree of crosslinguistic orthographic similarity is often calculated using the orthographic similarity metric (OS), an algorithm that was adapted from the graphic similarity index (GS), which was developed by Weber (1970). The original GS index, which was devised as an attempt to explain the degree to which reading substitution errors approximated participants' correct responses in terms of letters, includes seven sub-indices to which Weber assigned distinctive weights based upon her own intuitions about the importance of cues purported to be used during word identification. Weber's GS index is expressed by the following algorithm and comprises the following indices, as Example 1 below shows.

Example 1. Weber's Grapheme Similarity (GS) Index

$$GS = 10\left(\frac{[50F + 30V + 10C)}{A}\right] + 5T + 27B + 18E)$$

A: average number of letters between two words.

B: 1 if the first letter of the two words is the same; otherwise, B = 0

C: number of single letters shared by the two words.

E: 1 if the last letter of the two words is the same; otherwise, E = 0

F: number of pairs of adjacent letters in the same order shared by the two words.

T: ratio of number of letters in the shorter word to the number of letters in the longer one.

V: number of pairs of adjacent letters in reverse order shared the two words.

The GS index between two words ranges from 0 [zero] (no common letters) to 1,400 (identical letters). However, the application of the formula by itself can yield results that are difficult to interpret because the formula is sensitive to word length. Specifically, not all word pairs comprised of identical words will have the same GS index (Guasch, Boada, Ferré, & Sánchez-Casas, 2013). For example, the GS index of a word like *trim* with itself is 975, whereas the GS index of a word like *grime* with itself is 1,000. In order to address Weber's GS index, van Orden (1987) developed an improved orthographic similarity algorithm that is based on Weber's GS index algorithm. Van Ordern's algorithm (OS) yields similarity values that vary between 0 (no shared letters – minimum orthographic similarity) and 1 (identical letters – maximum orthographic similarity) and are not sensitive to word length(Guasch et al., 2013). Van Orden's algorithm is expressed below, where *P* is the first word of a pair and *R* the second word, as in Example 2 below.

Example 2. Van Ordern's Orthographic Similarity (OS) Algorithm

$$OS = \frac{GS(P,R)}{GS(R,R)}$$

Van Orden's algorithm, however, lacks an important property of similarity: the commutative property. Specifically, when comparing the words in a word pair, reversing the order in which they are input in the algorithm yields completely different orthographic similarity indices. With this drawback in mind, more recent studies have started to utilize the Levenshtein distance (Levenshtein, 1966) as a metric of orthographic similarity. Simply put, the Levenshtein distance between two words is the minimum number of single-character insertions, deletions or substitutions required to change one word into the other. Still, the metric is highly sensitive to word length. In order to neutralize those effects, Schepens et al. (2013) devised the normalized Levenshtein distance (NLD), which normalizes the metric and yields similarity scores that always range between 0 [zero] (when word pairs are orthographically dissimilar) and 1 (when word pairs are orthographically identical) irrespective of word length (Guasch et al., 2013). The NLD is calculated according to the algorithm shown in Example 3 below.

Example 3. Normalized Levenshtein Distance (NLD)

$$NLD = 1 - \frac{Levenshtein\ distance\ (string\ 1, string\ 2)}{Maximum\ string\ length\ (string\ 1, string\ 2)}$$

Since one of the major goals of the current study is to further understand how orthographic similarity between Portuguese-English cognates and false friends can mirror the organization of the bilingual mental lexicon, needless to say, choosing the most reliable as well as the most up-to-date metric of orthographic similarity was paramount. For this reason, the NLD seemed the most apposite orthographic similarity metric to use. To that end, NIM (Guasch et al., 2013), a web-based software developed by psycholinguists in the Research Center for Behavior Assessment at *Universitat Rovira iVirgili* in Tarragona, Spain, was utilized to calculate the NLD between the word pairs in the preliminary word list. One of the greatest features of NIM is the fact that it calculates both Levenshtein distance and NLD as well as Weber's GS and van Orden's OS indices. In addition, researchers can copy and paste lists of word pairs directly into NIM's interface, and the software will output a table with all the orthographic similarity indices, which can also be downloaded in Microsoft Excel format. Once the orthographic similarity values were obtained for all the word pairs in both word lists, they were entered into the master Excel spreadsheet according to their associated word pairs.

3.5. Obtaining Crosslinguistic Phonological Similarity Estimates

Phonological similarity (or distance), despite lacking a formally agreed upon definition in the literature, is often reduced to segmental similarity and explained in terms of distinctive features (Chomsky & Halle, 1992; Clements & Hume, 1995; G. N. Clements, 1985, *inter alia*) and other formal primitives of phonological theory. Assessing crosslinguistic phonological similarity via this traditional framework poses numerous challenges, in particular, because native and non-native speakers' mechanisms involved in speech perception and production are more multidimensional, taking into account not only segmental features but also prosodic and phonotactic properties (Bradlow et al., 2010; Mielke, 2012).

Because there currently is no standard phonological similarity metric readily available, many influential studies on the bilingual mental lexicon have often relied on subjective measures of phonological similarity, i.e., authors' and consultants' personal judgments as both native and non-native speakers, for crosslinguistic stimuli selection (cf., Comesaña et al., 2012; Friesen & Jared, 2012; Schwartz & Kroll, 2006; Schwartz, Kroll, & Diaz, 2007; *inter alia*). The present study avoided this shortcoming by computing an objective measure of crosslinguistic phonological similarity using a technique previously employed by Mielke (2012), viz., dynamic time warping (DTW), which was originally conceived to compare different speech patterns in automatic speech recognition and measures acoustic similarity between two time-dependent sequences, such as sound waveforms. In addition, because the likelihood that two words across different languages are cognates is highly correlated with their phonological similarity (Kondrak & Sherif, 2006), it is of utmost importance to this study to employ an objective metric such as DTW to measure phonological similarity.

DTW in its essence is a time sequence alignment algorithm that, for the purposes of the current study, iteratively aligns two wave files in a distance matrix by warping them at their beginning as well as at their end until an optimal match between the two wave files is found (Shinde & Pawar, 2014). To that end, the optimal path between two acoustically similar wave files is a diagonal line. The more the optimal path deviates from a diagonal line, the less acoustically similar two wave files are. *Figure 8* below illustrates the dynamically time warped the waveforms of the Portuguese-English cognate pair *grafite-graphite* and the false friend pair *smoking* 'tuxedo'-*smoking* recorded by native speakers of the respective languages. For comparison purposes, the dynamically time warped waveforms of the Portuguese word *grafite* with itself and the English word *smoking* with itself are also provided. Although the optimal path between the cognate pair and the false friend pair is not a straight diagonal line, it deviates less from a diagonal line in the false friend pair *smoking-smoking*, since the pronunciation of *smoking* in Portuguese is more acoustically similar to the English pronunciation of *smoking*. The optimal path for the cognate pair *grafite-graphite* starts rather jagged and completely deviates from a diagonal to a horizontal line after the fricative [f], indicating that the pronunciation of the Portuguese word *grafite* is less acoustically similar to the pronunciation the English word *smoking.* When Portuguese *grafite* and English *smoking* are compared with themselves, the optimal path is a straight diagonal line.

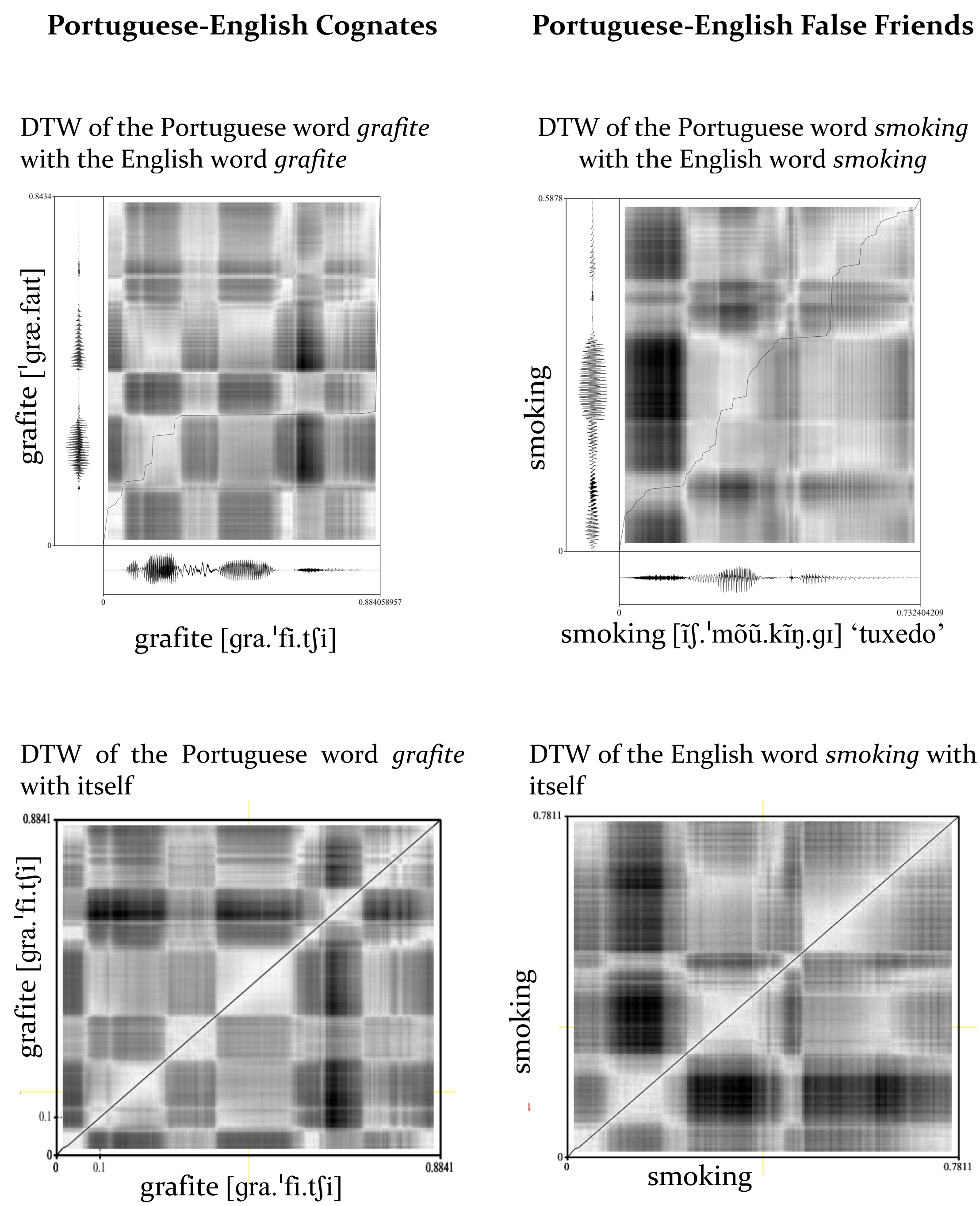


Figure 8: Dynamic time warping illustration of the English-Portuguese cognate pair grafite-graphite and the false friend pair smoking ‘tuxedo’-smoking. Areas that are darker represent points where the accoustic distance between the waveforms is less (greater acoustic similarity).

The DTW algorithm is usually implemented via software packages that quantify the degree to which waveforms of linguistic objects (sounds or words) are similar to each other. For the purposes of this study, *Phonological Corpus Tools* (PCT) (Hall, Allen, Fry, Mackie, & McAuliffe, 2016) was chosen for its easy-to-use interface, in particular, for its ability to perform all the acoustic-similarity-related computations unaided once a tab-delimited file that lists all the pairwise comparisons of individual waveforms to be compared is chosen. In addition, PCT also allows users to customize the acoustic analysis according to specific parameters related to the acoustic properties of a sound. The first of these customization parameters is *frequency limits* (or range), where researchers can set the minimum and the maximum frequency range (in Hz) for the acoustic analysis. According to Johnson (2011), human sensitivity to frequencies above 10,000 Hz is rather limited. Thus, "frequency components above 10,000 Hz are not likely to be useful for speech communication even if the speaker has perfect hearing" (p. 23). In light of this fact, this study set the limits of the frequency range to 80Hz minimum (to filter out low background noise) and 10,000 Hz maximum. The second customization parameter is *frequency resolution*, where researchers can set the number of filters to be used to divide up the frequency range they previously specified. The default setting of 26 filters was chosen for this study. The third parameter is the number of *Mel-frequency cepstral coefficients (MFCCs)*, where researchers can set the number MFCCs they desire, which are a type of representation of a sound that is based on the human ear response to sounds and are extensively used in automatic speech and speaker recognition systems. The default setting of 12 coefficients was chosen, which is consistent with the methodology in Mielke (2012). The fourth customization parameter is whether PCT should output the acoustic similarity results in a scale that ranges between 0 [zero] and 1, or use the default inverse similarity scale, which was chosen for this study. Finally, researchers can request that PCT generates a results table in the form of a tab-delimited text file, which can be saved on to the computer and later easily imported into Excel. *Figure* 9 below is a screen capture of the several parameters users can set in PCT. *Figure* 10 illustrates the results table PCT generates once it completes all the acoustic similarity computations in the pairwise comparison list.

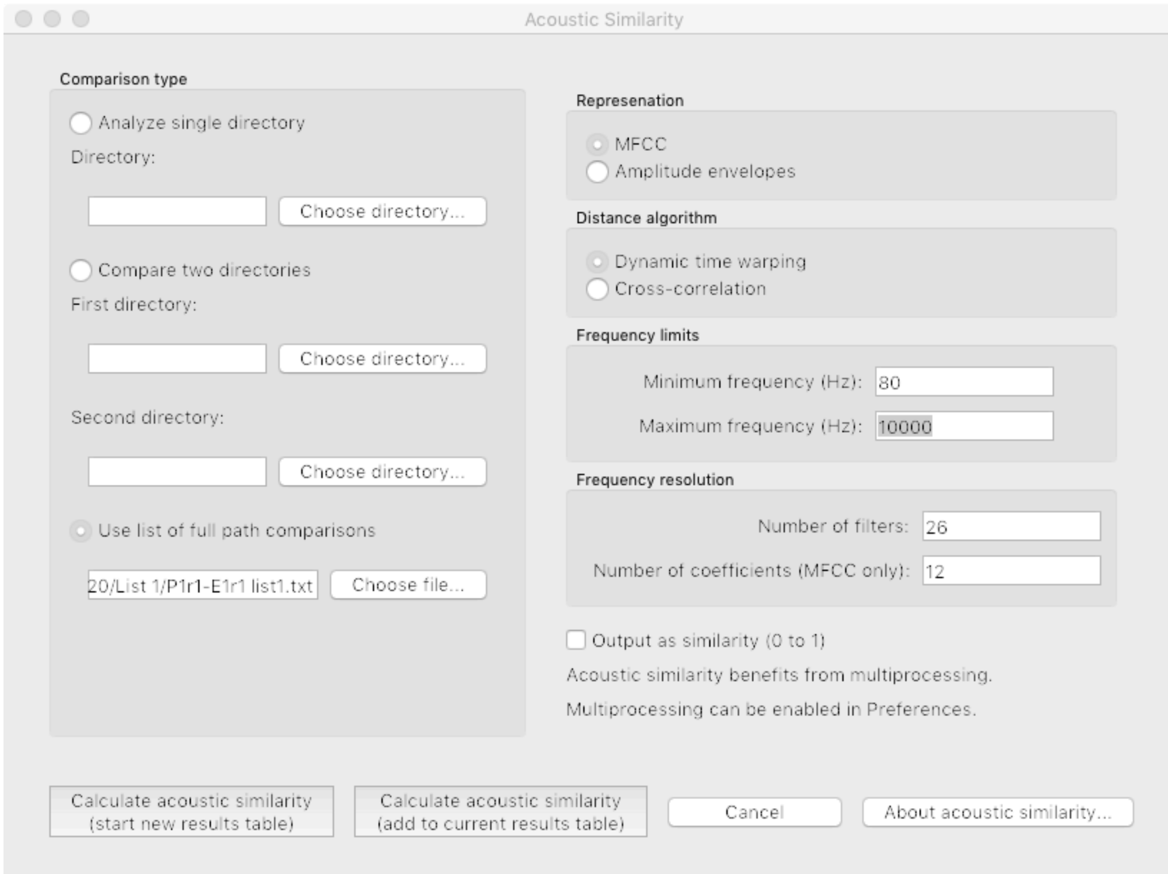


Figure 9. PCT screen capture illustrating the parameters users can set for DTW algorithm to be taken into account during the acoustic analysis

Acoustic similarity results

File 1	File 2	Representation	Match function	Minimum frequency	Maximum frequency	Result	Number of filters	Number of coefficients	Is similarity
AVG	AVG	MFCC	Dynamic time warping	80.0	10000.0	57.148	26	12	No
abate-1.wav	press.wav	MFCC	Dynamic time warping	80.0	10000.0	61.494	26	12	No
abrigo-1.wav	adoptive.wav	MFCC	Dynamic time warping	80.0	10000.0	61.417	26	12	No
abrupto-1.wav	abrupt.wav	MFCC	Dynamic time warping	80.0	10000.0	43.562	26	12	No
acaraje-1.wav	alarm.wav	MFCC	Dynamic time warping	80.0	10000.0	63.481	26	12	No
acento-1.wav	accent.wav	MFCC	Dynamic time warping	80.0	10000.0	48.264	26	12	No
acude-1.wav	alcohol.wav	MFCC	Dynamic time warping	80.0	10000.0	66.389	26	12	No
adepto-1.wav	adept.wav	MFCC	Dynamic time warping	80.0	10000.0	45.142	26	12	No
adesivo-1.wav	alignment.wav	MFCC	Dynamic time warping	80.0	10000.0	60.519	26	12	No
adjunto-1.wav	adjunct.wav	MFCC	Dynamic time warping	80.0	10000.0	42.859	26	12	No
adotivo-1.wav	adoptive.wav	MFCC	Dynamic time warping	80.0	10000.0	46.441	26	12	No
agenda-1.wav	agenda.wav	MFCC	Dynamic time warping	80.0	10000.0	48.624	26	12	No
agulha-1.wav	angle.wav	MFCC	Dynamic time warping	80.0	10000.0	61.454	26	12	No
aipo-1.wav	signature.wav	MFCC	Dynamic time warping	80.0	10000.0	76.781	26	12	No
alagamento-1.wav	astute.wav	MFCC	Dynamic time warping	80.0	10000.0	65.394	26	12	No
alcatrao-1.wav	baggage.wav	MFCC	Dynamic time warping	80.0	10000.0	66.609	26	12	No
alcool-1.wav	alcohol.wav	MFCC	Dynamic time warping	80.0	10000.0	52.642	26	12	No
alecrim-1.wav	band.wav	MFCC	Dynamic time warping	80.0	10000.0	59.693	26	12	No
alho-1.wav	rim.wav	MFCC	Dynamic time warping	80.0	10000.0	67.209	26	12	No
alvenaria-1.wav	base.wav	MFCC	Dynamic time warping	80.0	10000.0	73.371	26	12	No
ameaca-1.wav	smoking.wav	MFCC	Dynamic time warping	80.0	10000.0	58.953	26	12	No
amigo-1.wav	battery.wav	MFCC	Dynamic time warping	80.0	10000.0	67.481	26	12	No
anao-1.wav	beige.wav	MFCC	Dynamic time warping	80.0	10000.0	70.293	26	12	No
anca-1.wav	bomb.wav	MFCC	Dynamic time warping	80.0	10000.0	59.523	26	12	No
angu-1.wav	brave.wav	MFCC	Dynamic time warping	80.0	10000.0	74.006	26	12	No
angulo-1.wav	angle.wav	MFCC	Dynamic time warping	80.0	10000.0	52.611	26	12	No
aplicacao-1.wav	application.wav	MFCC	Dynamic time warping	80.0	10000.0	50.841	26	12	No
apoio-1.wav	capacity.wav	MFCC	Dynamic time warping	80.0	10000.0	61.988	26	12	No
arnica-1.wav	cargo.wav	MFCC	Dynamic time warping	80.0	10000.0	59.885	26	12	No
assiste-1.wav	assist.wav	MFCC	Dynamic time warping	80.0	10000.0	53.071	26	12	No
audiencia-1.wav	audience.wav	MFCC	Dynamic time warping	80.0	10000.0	55.489	26	12	No
avestruz-1.wav	car.wav	MFCC	Dynamic time warping	80.0	10000.0	75.978	26	12	No
baba-1.wav	cheque.wav	MFCC	Dynamic time warping	80.0	10000.0	71.889	26	12	No
baderna-1.wav	chocolate.wav	MFCC	Dynamic time warping	80.0	10000.0	57.238	26	12	No
banco-1.wav	bank.wav	MFCC	Dynamic time warping	80.0	10000.0	49.713	26	12	No
bando-1.wav	coma.wav	MFCC	Dynamic time warping	80.0	10000.0	60.396	26	12	No
baque-1.wav	comatose.wav	MFCC	Dynamic time warping	80.0	10000.0	62.667	26	12	No
barata-1.wav	comedy.wav	MFCC	Dynamic time warping	80.0	10000.0	60.564	26	12	No

Reopen function dialog | Save to file | Close window

Figure 10. Acoustic similarity results table. In the Result column, higher values represent greater acoustic distances (less acoustic similarity) between the waveforms.

To obtain the waveforms for each Portuguese and English word in the official word lists, two male native speakers of each language of interest were recorded. To accomplish this, one native speaker of American English (midwestern accent, native of Ohio) and one native speaker of Canadian English (southeastern Ontario accent, native of Cornwall) were recorded reading aloud each English word. Conversely, two native speakers of Brazilian Portuguese, one native of São Paulo and one native of Rio de Janeiro, who were born and raised in the metro areas of both cities, were recorded reading aloud each Portuguese word. For the recording session, two spreadsheets were created (one for each language) and then converted to individual slides in a PDF file. Each slide had a white background with one word in a legibly large font in its center. Recordings took place in the Language Acquisition Research Lab at the University of Ottawa. Readers sat at a desk in front of a computer screen and were accompanied by the researcher. They were instructed to read aloud each of the words that appeared in the center of the computer screen as naturally as possible. The researcher controlled the pace in which each word was displayed. Each reader was recorded twice, with 15-minute break in between recordings to control for pronunciation discrepancies. Audacity® ("Audacity®: Free Audio Editor and Recorder," 2019) was selected for the recordings. Audio was captured using a condenser microphone connected to a preamplifier and saved into a waveform file (.wav) at 44,100 Hz. Audacity® was also used to create a label track that contained the orthographic representation of each word in the recording. Each labeled word was then exported as a separate waveform file, with its orthographic representation for its name, and saved according to reader's name and recording session. Both official word lists (A + B) were used to create tab-delimited text files that listed all the pairwise comparisons in the list according to their specific file location on the computer, which were then used in PCT to perform the required acoustic similarity computations.

In order for the phonological similarity metric to be as accurate as possible, it was necessary to account for readers' dialectal differences in pronunciation as well as any pronunciation discrepancies that might have occurred between the first and second recording (before and after the break). To that end, each recording from each reader of Portuguese was compared with each recording from each reader of English. In the end, eight sets acoustic similarity comparisons (see Table 6 below) were performed using PCT. Once all the DTW acoustic similarity values were obtained, they were entered into an Excel spreadsheet so that the mean crosslinguistic phonological similarity overlap could be calculated for the data analysis.

Table 6. Acoustic similarity computations performed in PCT

<table>
<tr><th>Reader's
Native Language and
Place of Origin</th><th colspan="2">Recording
r1 – before break
r2 – after break</th><th>Acoustic similarity
calculations performed in
PCT</th></tr>
<tr><td>P1 – Portuguese
Rio de Janeiro</td><td>r1</td><td>r2</td><td rowspan="4">• P1r1-E1r1
• P1r1-E1r2
• P1r2-E1r1
• P1r2-E1r2
• P2r1-E2R1
• P2r1-E2r2
• P2r2-E2r1
• P2r2-E2r2</td></tr>
<tr><td>P2 – Portuguese
São Paulo</td><td>r1</td><td>r2</td></tr>
<tr><td>E1 – English
Ohio</td><td>r1</td><td>r2</td></tr>
<tr><td>E2 – English
Ontario</td><td>r1</td><td>r2</td></tr>
</table>

3.6. Experimental Procedure

The experimental task consisted of a lexical decision task with masked priming, which was solely conducted in Presentation® ("Presentation®: Precise, Powerful Stimulus Delivery," 2017), a stimulus delivery and experiment control software widely used in psycholinguistic experiments. Data collection occurred over a period of ten months in the Language Acquisition Research Laboratory at the University of Ottawa. Participants sat at a large computer desk in front of a 23-inch LCD monitor. Before participants were allowed to initiate the experimental task, they engaged in a prescreening interview with the researcher, in which they were asked a few questions regarding their personal background and level of education, as well as how they had learned about the experiment and whether they had any particular questions or concerns. Participants who were native speakers of Portuguese were interviewed in Portuguese. Those who were speakers of English speakers were interviewed in English. The purpose of the prescreening interview was to ensure participants met the language requirements of the study and to give them a chance to get themselves situated in the laboratory. Following the prescreening interview, participants were provided with the consent form (see Appendix C) and were given a few minutes to read and sign it. Participants were then assigned a pseudocode that consisted of a sequence of random letters and numbers so that their data would remain anonymous and were instructed to use their assigned pseudocode thereafter. Subsequently, participants were asked to complete the online Language Background Questionnaire (introduced in 3.1., *Google Form*s, see Appendix D) on the computer as well as the Language Proficiency Test (cloze), which was administered on paper. Most participants completed these steps in approximately 20-30 minutes.

Afterwards, the experimenter started Presentation®, and the software randomly assigned participants to one of the two word lists as well as instructed them to enter their assigned pseudocode in a text box that appeared on the computer screen. Presentation® would then prompt participants to press the blue button on a *Cedrus RB-730 Response Pad* (button box) to indicate they were ready commence the experimental task. The buttons on the button box had four distinct colors (blue, green, red and white), and the ones participants would be pressing during the task (blue, green and red) had a label in a legible font indicating their specific function in the experimental task. Once participants pressed the blue button, they were presented with the task instructions and instructed to press the blue button again to begin the (20) practice trials. The purpose of the practice trials was to ensure participants had understood the task instructions and become familiar with the procedure. Once participants completed the practice trials, the researcher checked whether they needed further clarification about the procedure. When participants were cleared to continue, they were instructed to press the blue button one more time to commence the experimental trials.

Each experimental trial consisted of a sequence of five events. First, a forward mask consisting of ten hash marks (##########) was presented in the center of the computer screen for 500ms. Second, the forward mask was followed by the Portuguese prime in lowercase letters for 60ms. Third, the English target was presented in uppercase letters for 500ms. Fourth, once the English target was displayed, participants had 3,000ms (3s) to decide by the press of a button (green = yes, red = no) whether or not the target was a real word of English. Fifth, as soon as participants pressed the button to indicate their response, the next prime-target pair was presented automatically. If participants did not press any button during the allotted 3,000ms, the next prime-target pair was presented.

The experimental task was comprised of a total of four blocks of 150 trials, of which four practice pairs were always displayed at the beginning of each block. A pause was added at the end of each block, and participants were encouraged to get up, stretch, rehydrate or use the restroom during this break. On average, participants completed the task within 45 minutes, including break time. Their reaction time and accuracy were recorded for the data analysis.

The following chapter presents the results of the study in detail.

4. The Study: Results

In order to facilitate the presentation of the results of the study, this chapter was divided into two sections, *Preliminary Analyses* and *Mixed-Effects Models*. In the *Preliminary Analyses* section, both dependent variables, reaction time and accuracy, and each of the independent variables that comprised the master dataset were independently analyzed, and whenever plotting the variable enhanced the results being discussed, plots were provided. The results provided in this section serve as groundwork for the ensuing section, since only more conventional statistical analysis tools, such as correlations, t-tests and ANOVAs, were used. In the *Mixed-Effects Models* section, one linear mixed-effects analysis of the relationship between reaction time and the independent variables (`lmer` model), as well as one logistic regression analysis between accuracy and the independent variables (`glmer` model) are presented. Both mixed-effects analyses allowed for precise quantification of the individual effects of crosslinguistic similarity, which was one of the primary goals of this study. Both SPSS ("IBM SPSS Statistics 25," 2018) and RStudio (RStudio Team, 2018), were used extensively to carry out the analyses in this section. In addition, both SPSS and `ggplot2` (Wickham, 2016), a data visualization package developed for the statistical programing language R, were used to facilitate the presentation of the results.

Before proceeding with the data analysis, it is important to underscore the fact that prime-target pairs that were originally coded as *Practice* did not enter the analyses. As discussed in chapter 3, 20 practice pairs were displayed to participants at the beginning of the experiment in order to give participants an opportunity to practice the experimental task and ask the researcher any questions they might have had about the procedure. Four practice pairs were also displayed at the beginning of a new block after each of the planned pauses/breaks.

4.1. Preliminary Analyses[7]

4.1.1. Dependent Variables

4.1.1.1. RT (Reaction Time/Response Latency)

As mentioned in Chapter 3, Presentation ("Presentation®: Precise, Powerful Stimulus Delivery," 2017), a stimulus delivery and experiment control software, was used to control the presentation of the stimuli and record participants' reaction time during the study. To that end, Presentation records reaction time in tenths of milliseconds. However, for the sake of consistency with psycholinguistic studies, which normally report reaction time in milliseconds, the variable RT was converted from tens of milliseconds to milliseconds. The descriptive statistics for RT, including a breakdown by language group (hereinafter referred to as the independent variable

[7] A brief description of the variable in the master dataset is provided hereinafter between parentheses.

LANGUAGE) are provided in Table 7 below. Kernel density plots (Figure 11) are also provided to facilitate visualization of the distribution of RT.

Table 7. RT (Descriptive Statistics)

	N	Mean	Std. Dev.	Minimum	Maximum
Overall	33,516	769.01	316.04	291.20	3,000
Bilinguals	16,464	808.60	348.29	291.20	3,000
Functional Monolinguals	17,052	730.79	276.08	305.60	3,000

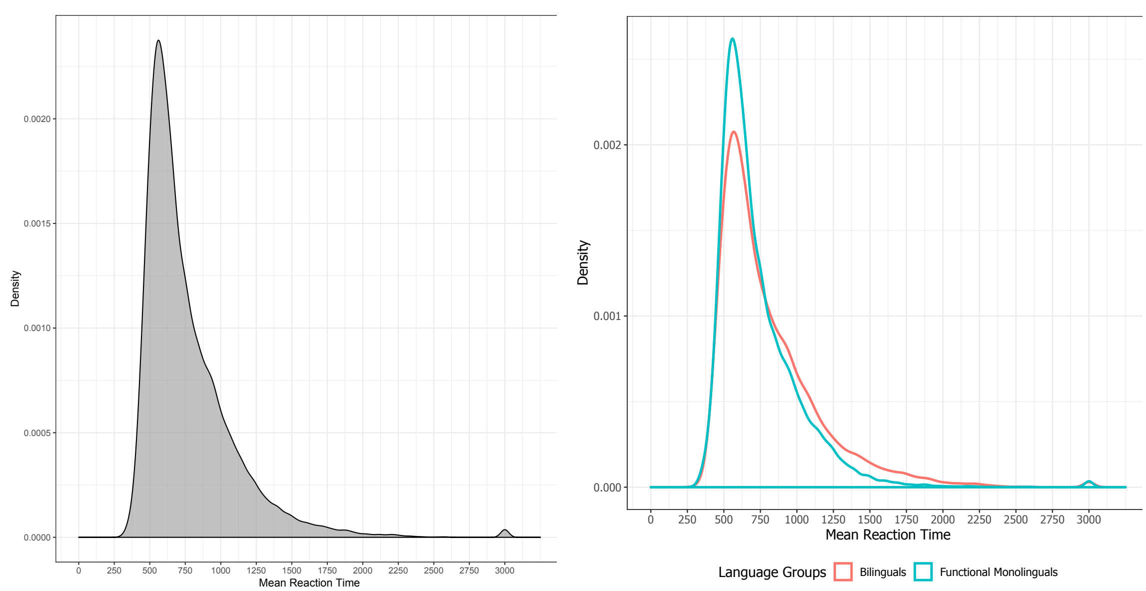


Figure 11. Kernel density plots illustrating the distribution of RT (overall and by LANGUAGE)

To further examine the differences in mean reaction time between bilinguals and functional monolinguals, an independent samples t-test was performed on RT with LANGUAGE as the grouping factor. The t-test showed that functional monolinguals were on average significantly faster (M = 730.79, SD = 276.08) than bilinguals (M = 808.60, SD = 348.29), $t(33{,}514) = -22.705$, $p < .001$. Whether or not this significant difference between bilinguals and functional monolinguals played a major role in mean reaction time when all other variables are jointly taken into account will be explored in the Mixed-Effects Models section.

4.1.1.2. RAW_ACCURACY (Response Accuracy)

Presentation was also used to record participant's accuracy. As described in Chapter 3, participants had 3,000ms in each trial (prime-target pair) to decide by the press of a button whether or not the target word presented was a genuine English word. Accuracy was thus recorded as a hit (correct answer), incorrect, or miss (when the allocated 3,000ms was exceeded) for each trial, making it a categorical variable. As Table 8 shows, irrespective of language group, the number of correct answers in the entire experiment, was higher than the number of incorrect answers. However, when the data is broken down by language group, bilinguals had a higher number of correct answers than functional monolinguals.

Table 8. RAW_ACCURACY (Count)

	Hit (correct)	Incorrect	Miss
Overall	30,956	2,472	88
Bilinguals	16,297	713	42
Functional Monolinguals	14,659	1,759	46

A point-biserial correlation between RT and RAW_ACCURACY was run to better understand the relationship between the two dependent variables. Since this type of correlation is run between a continuous and a dichotomous variable, only correct and incorrect answers were considered. The correlation yielded a weak yet significant negative correlation between RT and RAW_ACCURACY, $r(33{,}428) = -.186$, $p < .001$, suggesting that on average as participants' accuracy increased, their reaction time decreased.

In order to determine the percentage of hits (correct answers) per participant and to weed out participants with an unusually low percentage of correct answers, a new variable (CORRECT_PERCENTAGE) was created, which in essence transformed RAW_ACCURACY from a categorical to a continuous variable. To accomplish this, the total number of hits (correct answers) per participant was divided by the total number of trials in the experiment. This result was then multiplied by 100%. Because the percentage of correct answers for one of the bilingual participants was significantly lower than bilingual participants' mean, that specific participant had to be removed from the analysis.

Table 9. CORRECT_PERCENTAGE (Descriptive Statistics)

Language Group	N	Mean	Std. Dev.	Min.	Max.
Combined	33,428	92.17	6.59	64.10	99.20
Bilinguals	16,418	88.72	7.43	64.10	98.60
Functional Monolinguals	17,010	95.50	3.08	84.80	99.20

An independent samples t-test was performed on CORRECT_PERCENTAGE with LANGUAGE as a grouping factor. The t-test showed that functional monolinguals' percentage of correct answers (M = 95.50, SD = 3.08) was significantly higher than bilinguals' (M = 88.72, SD = 7.43), t(33,426) = 109.54, p < .001. Although this result is only preliminary, it shows that functional monolinguals were on average more accurate at deciding whether the presented word was a genuine word of English.

4.1.2. Independent Variables

4.1.2.1. SEMAC and ERS (English Proficiency)

As discussed in Chapter 3, English proficiency was assessed via a cloze proficiency test. Participants' answers on the test were scored using both the semantically acceptable scoring procedure (SEMAC) and the exact replacement scoring method (ERS), producing two language proficiency variables, SEMAC and ERS, which, as indicated by previous research, are positively correlated, $r(33{,}516) = .839$, $p < .001$. An independent samples t-test was performed on both ERS and SEMAC with LANGUAGE as the grouping factor. With respect to SEMAC, the t-test revealed that functional monolinguals were significantly more proficient in English than bilinguals, $t(33{,}514) = 119.97$, $p < .001$. Similarly, with respect to ERS, functional monolinguals were also shown to be more proficient in English than bilinguals, $t(33{,}514) = 116.88$, $p < .001$. These results show that irrespective of scoring method functional monolingual participants were on average significantly more proficient in English than bilinguals, which was expected since this group of participants was comprised of native speakers of the language. Please refer to Table 10 below for the mean differences between both variables.

Table 10. ERS and SEMAC Independent samples t-test (Descriptive Statistics)

	Language Group	N	Mean	Std. Dev	Mean Diff.
ERS	Bilinguals	16,464	50.07	11.75	13.51*
	Functional Monolinguals	17,052	63.59	9.31	
SEMAC	Bilinguals	16,464	76.86	16.28	15.69*
	Functional Monolinguals	17,052	92.55	5.09	

*$p < .001$

In order to determine the relationship between both ERS and SEMAC with RT, irrespective of language group, a Pearson product-moment correlation was run. A weak yet significant negative correlation was obtained for both ERS, $r(33{,}516) = -.10$, $p < .001$, and SEMAC, $r(33{,}516) = -.13$, $p < .001$. These results indicate that on average as mean English proficiency increased, mean reaction time decreased irrespective of language group.

The relationship between English proficiency and RT presents itself differently when the data is broken down by LANGUAGE. For bilinguals, a weak negative correlation is maintained with both ERS and SEMAC (both *p*s <.001). For functional monolinguals, a weak positive correlation was found with ERS ($p < .001$). However, no significant correlation was found with SEMAC. See Table 11 below for the results of the tests.

Table 11. *Pearson product-moment correlations between RT and ERS and SEMAC*

Language Group	**Scoring Method**	***r***
Bilinguals	ERS	r(16,464) = – .10*
	SEMAC	r(16,464) = – .11*
Functional Monolinguals	ERS	r(17,052) = .04*
	SEMAC	r(17,052) = 0

* $p < .001$

A point-biserial correlation was run in order to evaluate the relationship between both ERS and SEMAC with RAW_ACCURACY. A significant positive correlation was found with both variables (both *p*s < .001). When the English proficiency data was split by LANGUAGE, significant positive correlations were also found, indicating that as mean English proficiency increased, so did mean accuracy. Table 12 below shows the results of the point-biserial correlations.

Table 12. Point-biserial correlations between ERS and SEMAC with RAW_ACCURACY

Language Group	**Scoring Method**	***r***
Bilinguals	ERS	r(16,418) = .11*
	SEMAC	r(16,418) = .12*
Functional Monolinguals	ERS	r(17,010) = .05*
	SEMAC	r(17,010) = .05*

Since the correlations between SEMAC scores and RT and RAW_ACCRURACY were stronger, SEMAC scores were selected to enter the mixed-model analyses as an indicator of participants' English proficiency level (PROFICIENCY).

4.1.2.2. SEMANTIC (Crosslinguistic Semantic Similarity/Overlap)

Crosslinguistic semantic similarity was coded in the master dataset as a categorical variable consisting of four categories according to the degree of crosslinguistic semantic overlap between a prime and its target: cognates (COG), false friends (FF), controls (CTRL), and distractors (DISTR). Consistent with the results discussed in Alves-Soares, (2013) and other previous linguistic experiments that have utilized masked priming in their methodology, participants' reaction time was highly influenced by the degree of semantic relationship between prime and target. Mean reaction time was lowest for targets that were matched to a prime with which they share a high degree of crosslinguistic semantic overlap (cognate items), and highest for pseudoword targets (DISTR items), i.e., when a random word of Portuguese was matched with a pseudoword target. This pattern is also present when the data is split by language groups, suggesting that a priming effect of crosslinguistic semantic similarity was present for both language groups. Splitting the data by LANGUAGE shows that functional monolingual participants were on average faster to decide whether or not the target was a word of English than bilinguals. Refer to Table 13 for the descriptive statistics.

A two-way ANOVA was then performed with RT as the dependent variable and LANGUAGE and SEMANTIC and as factors to further investigate these initial observations. The ANOVA revealed a significant main effect of semantic similarity, $F(3, 33{,}508) = 875.88$, $p < .001$, a significant main effect of language group, $F(1, 33{,}508)$

= 333.88, $p < .001$, as well as a significant interaction between semantic similarity and language group, $F(3, 33{,}508) = 10.06$, $p < .001$. Pairwise comparisons using the Bonferroni adjustment indicated that functional monolinguals were on average 73ms faster than bilinguals ($p < .001$). Regarding reaction time differences due to semantic similarity, in comparison to controls, participants from both language groups were on average 45ms faster to respond to cognates and 194ms slower to respond to distractors (both *p*s < .001). Participants were also on average 29ms faster to respond to cognates than false friends ($p < .001$). No significant mean differences in reaction time were observed between control and false friends ($p = .079$).

Table 13. RT by SEMANTIC: Descriptive Statistics

Language Group	Semantic Similarity	N	Mean*	Std. Dev.*	Min.*	Max.*
Combined	COG	5,489	657.99	261.42	291.20	3,000
	FF	2,890	687.34	285.97	326.00	3,000
	CTRL	8,379	703.50	280.18	339.70	3,000
	DIST	16,758	852.22	331.32	313.4	3,000
Bilinguals	COG	2,700	680.35	290.99	291.20	3,000
	FF	1,416	733.33	332.78	341.00	3,000
	CTRL	4,116	737.77	308.32	345.80	3,000
	DIST	8,232	899.03	362.28	332.90	3,000
Functional Monolinguals	COG	2,789	636.35	227.17	305.60	3,000
	FF	1,474	643.17	223.68	326.00	3,000
	CTRL	4,263	670.42	245.58	339.70	3,000
	DIST	8,526	807.02	291.37	313.40	3,000

* in milliseconds

As SEMANTIC, LANGUAGE, and ACCURACY_RAW are all categorical variables, in order to examine the effect of semantic similarity on accuracy by language group, a breakdown of total number of hits and incorrect answers as by language groups and semantic similarity was obtained (see Table 14). The percentage of correct and incorrect answers was also entered into the breakdown so that any trends could be more easily spotted.

Table 14. RAW_ACCURACY by SEMANTIC and LANGUAGE

Language Group			Raw Count	Percentage (%)
Bilinguals	COG	incorrect	79	2.9
		hit	2,615	97.1
		Total	2,694	100.0
	FF	incorrect	107	7.6
		hit	1,305	92.4
		Total	1,412	100.0
	CTRL	incorrect	256	6.2
		hit	3,849	93.8
		Total	4,105	100.0
	DISTR	incorrect	1,317	16.0
		hit	6,890	84.0
		Total	8,207	100.0
Functional Monolinguals	COG	incorrect	45	1.6
		hit	2,739	98.4
		Total	2,784	100.0
	FF	incorrect	32	2.2
		hit	1,441	97.8
		Total	1,473	100.0
	CTRL	incorrect	111	2.6
		hit	4,138	97.4
		Total	4,249	100.0
	DISTR	incorrect	525	6.2
		hit	7,979	93.8
		Total	8,504	100.0

A particular trend clearly indicated by Table 14 is the fact that functional monolinguals' percentage of correct answers is higher than bilinguals' across all semantic categories. In addition, distractor pairs yielded the lowest percentage of correct answers for both language groups while cognate pairs yielded the highest. Together these two trends indicate a possible main effect of LANGUAGE and SEMANTIC on RAW_ACCURACY.

4.1.2.3. NLD (Crosslinguistic Orthographic Similarity/Overlap)

As discussed in Chapter 3, to measure crosslinguistic orthographic similarity (overlap) between primes and targets that were selected for the experiment, the Normalized Levenshtein Distance (NLD) (Schepens et al., 2013) was selected. The NLD is the most up-to-date and reliable metric for the purposes of this study. It yields similarity values that range between 0 [zero] (when prime-target pairs are orthographically dissimilar) and 1 (when prime-target pairs are orthographically identical), irrespective of word length (Guasch et al., 2013). Table 15 below provides the descriptive statistics of the variable NLD by SEMANTIC. The degree of crosslinguistic orthographic overlap between prime-target pairs classified as COG and FF is significantly higher than in pairs classified as CTRL and DISTR. The mean differences can be more clearly seen in both Table 15 and Figure 12 below.

Table 15. NLD: Descriptive Statistics

Semantic Similarity	**N**	***Mean***	***Std. Dev***	***Min.***	***Max.***
COG	5,489	.71	.12	0	1
FF	2,890	.70	.20	0	1
CTRL	8,379	.12	.12	0	1
DISTR	16,758	.11	.12	0	1

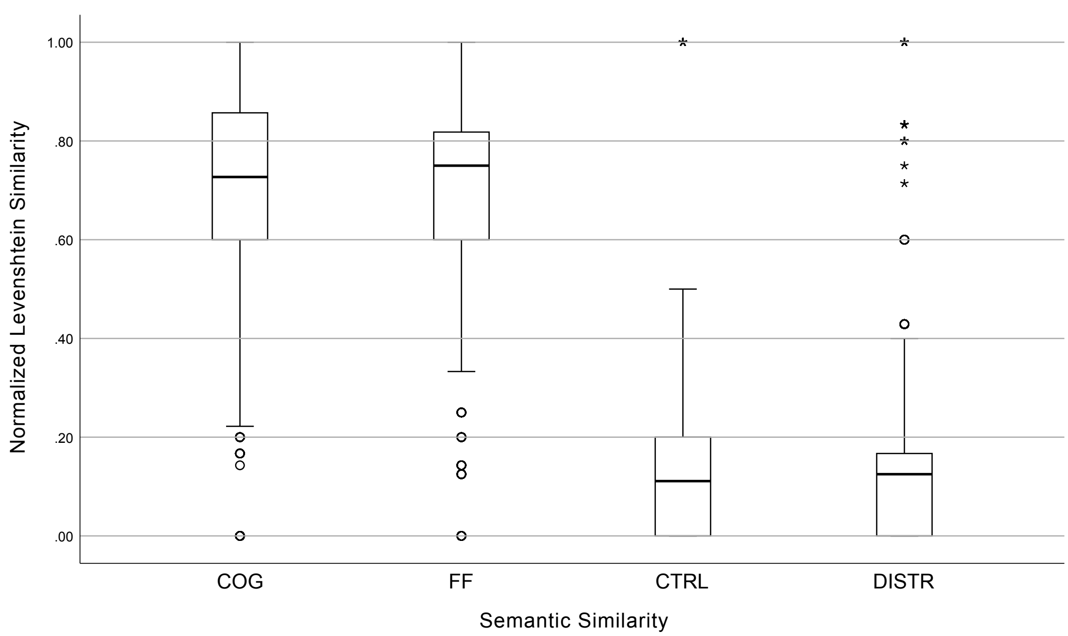


Figure 12. Mean differences in crosslinguistic orthographic overlap across semantic categories showing a higher mean NLD for COG and FF pairs.

4.1.2.4. PHONETIC (Crosslinguistic Phonological Similarity/Overlap)

As discussed in Chapter 3, the present study computed an objective crosslinguistic phonological similarity metric using a technique previously employed by Mielke (2012), viz., dynamic time warping (DTW). To that end, recordings of two native speakers of Portuguese and of two native speakers of English reading aloud the selected experimental words in their native language were made and used to compute a mean crosslinguistic phonological similarity value between prime-target pairs. Lower crosslinguistic phonological similarity values represent a greater degree of crosslinguistic phonological overlap between a prime and its target.

One hurdle that became apparent during the preliminary stages of the data analysis pertained to distractor pairs, because no crosslinguistic phonological similarity value for them could have been computed via the DTW algorithm. As prefaced in Chapter 3, distractor pairs in the experiment consisted of a genuine word of Portuguese as a prime that was matched with a pseudoword as a target. Even though the pseudowords selected for the experiment were generated by a pseudoword generator application that had been programed to generate pseudowords that resembled words of English as closely as possible, some of the consonant clusters as well as consonant-vowel combinations were not easily pronounceable by a native speaker of English due to the phonotactics of the language. As a result, recording a native speaker of English reading the list of pseudowords aloud as naturally as possible was a major challenge. While programing a speech synthesizer and requesting Siri to speak the pseudowords to the make the recordings were two options explored, the end result sounded quite artificial and robotic, lacking the connected and coarticulatory features present in human speech. Moreover, leaving the cells pertaining to distractor pairs blank in the crosslinguistic phonological similarity data would certainly lead to convergence complications in the later stages of the data analysis, particularly for fitting linear mixed-effects models. Thus, it was paramount that the distractor pairs cells be populated with data.

The most elegant solution reached was to populate the cells with data that followed the same distribution of the DTW crosslinguistic phonological similarity data. To achieve this, first the crosslinguistic phonological similarity data was plotted to ensure it followed a Laplace-Gaussian (normal) distribution and its descriptive statistics were obtained. With this information available, using the `rnorm` function in R, random numbers that followed a normal distribution with the exact same means and standard deviation of the original crosslinguistic phonological similarity data were generated. Afterward, both the original data and the randomly generated data were consolidated in a new variable, PHONETIC. Finally, a Pearson product-moment correlation was carried out and descriptive statistics were obtained to ensure no disparities had arisen from the procedure. Table 16 and Figure 13 below show the descriptive statistics and the kernel density plots of the crosslinguistic phonological similarity data before and after the procedure. To more easily illustrate the mean differences across the four semantic categories, descriptive statistics and box plots of the data were also obtained for the variable PHONETIC (see Table 17 and Figure 14).

Table 16. Original DTW crosslinguistic phonological similarity data vs. consolidated data

Original Data					Consolidated Data				
N	Mean	SD	Min.	Max.	N	Mean	SD	Min.	Max.
16,758	57.07	5.56	45.50	85.88	33,516	57.07	5.56	35.42	85.87
r(33,516) = 1, p < .001									

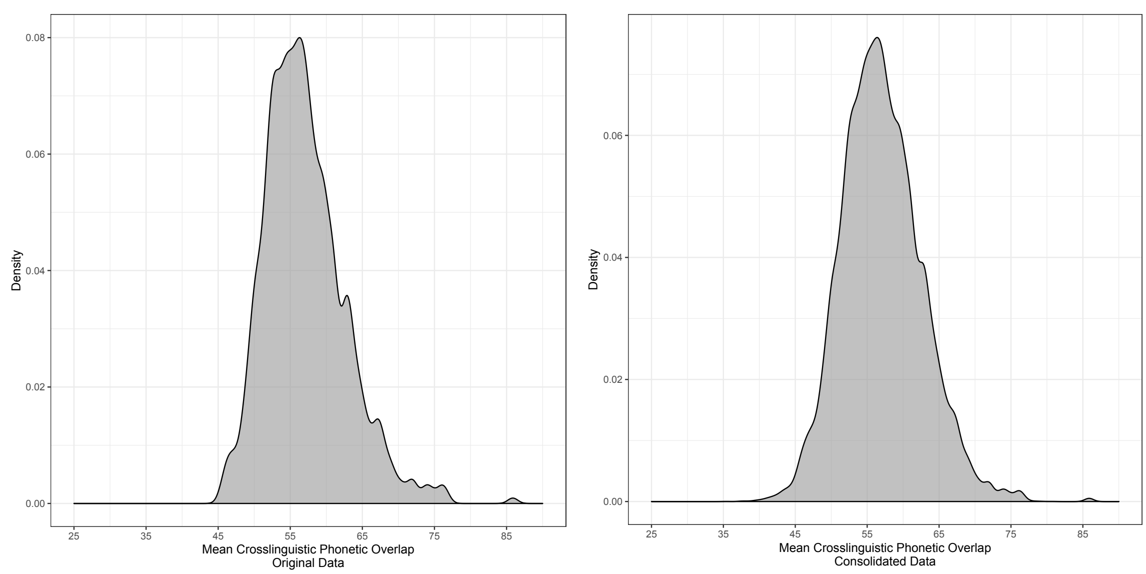


Figure 13. Kernel density plots of the DTW crosslinguistic phonological similarity data before (left) and after (right) consolidation.

Table 17. PHONETIC: Descriptive Statistics

Semantic Similarity	**N**	***Mean***	***Std. Dev***	***Min.***	***Max.***
COG	5,489	54.32	4.35	45.60	71.82
FF	2,890	53.74	3.44	45.51	67.98
CTRL	8,379	60.03	5.26	48.62	85.87
DISTR	16,758	57.06	5.55	35.42	80.74

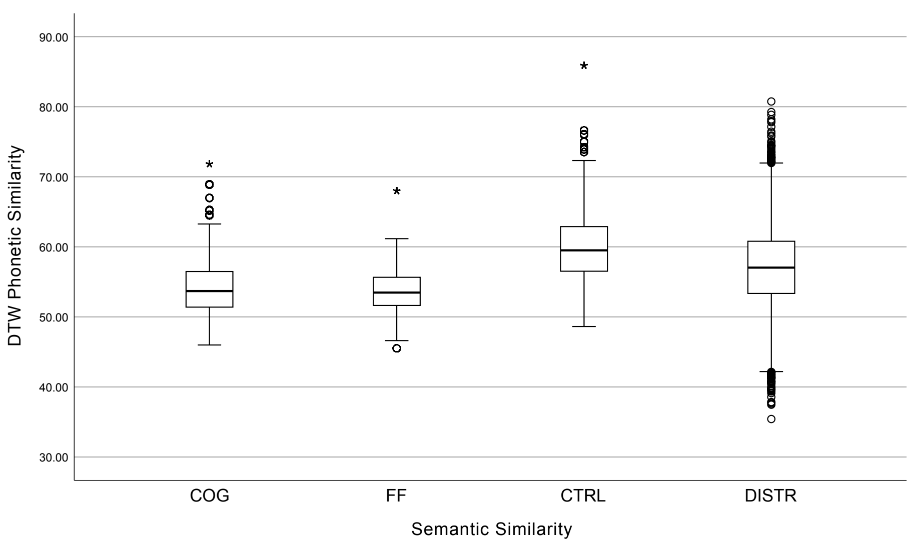


Figure 14. Mean differences in DTW crosslinguistic phonological overlap across semantic categories showing a lower mean for COG and FF pairs. Lower values represent a greater the degree of crosslinguistic phonological overlap between a prime and its target.

With the transformations on the DTW crosslinguistic phonological similarity data complete, the relationship between the PHONETIC with both the two dependent variables could be evaluated. A Pearson product-moment correlation between PHONETIC and RT yielded a weak yet significant positive correlation between the two variables, $r(33{,}516) = .02$, $p = .002$, suggesting that on average participants reacted more slowly to more phonologically dissimilar pairs. A point biserial correlation between PHONETIC and RAW_ACCURACY did not yield a statistically significant result, $r(33{,}428) = -.006$, $p = .31$.

4.1.2.5. English_Freq and Portuguese_Freq (Word Frequency)

As discussed in chapter 3, a word frequency count, viz., how often a particular word occurs per million words in collected corpora was obtained for each Portuguese and English word selected for the present study. To that end, *O corpus do Português* (Davies, 2017b), a large online corpus of Portuguese words, was used to obtain word frequency count for the Portuguese words, while *Corpus of Contemporary American English (COCA)* (Davies, 2017a), was used to obtain word frequency count for the English words.

Table 18 below shows the descriptive statistics for both English_Freq and Portuguese_Freq by SEMANTIC. It is important to underscore the fact that COCA yielded a word frequency count of zero for all of the pseudoword targets in distractor pairs selected for the experiment, since these targets are not words that exist in the

English language. Because of this, distractor pairs have zero mean English frequency in Table 18 below.

Table 18. *English_Freq and Portuguese_Freq by SEMANTIC (Descriptive Statistics)*

	SEMANTIC	N	Mean	Std. Dev.	Min.	Max.
COG	English_Freq	5,489	31.77	76.76	0.14	735.39
	Portuguese_Freq	5,489	36.58	84.74	0.03	688.23
FF	English_Freq	2,890	36.15	94.06	0.08	774.33
	Portuguese_Freq	2,890	36.28	83.68	0.00	562.51
CTRL	English_Freq	8,379	33.77	84.83	0.08	774.33
	Portuguese_Freq	8,379	18.50	50.58	0.00	443.81
DISTR	English_Freq	16,758	0.00	0.00	0.00	0.00
	Portuguese_Freq	16,758	19.48	66.77	0.00	688.23

A Pearson product-moment correlation was run to examine the relationship between English_Freq and Portuguese_Freq as well as between both word frequency variables and RT. The results of the correlation between English_Freq and Portuguese_Freq indicated positive relationship between the two variables, $r(33{,}516) = .15$, $p < .001$. The correlation with RT indicated a weak yet significant negative relationship with both English_Freq, $r(33{,}516) = -.11$, $p < .001$, and Portuguese_Freq, $r(33{,}516) = -.04$, $p < .001$. A point biserial correlation between both word frequency variables and RAW_ACCURACY indicated a weak yet significant positive correlation for both English_Freq, $r(33{,}428) = .06$, $p < .001$, and Portuguese_Freq, $r(33{,}428) = .02$, $p < .001$.

4.1.2.6. Time

Time as an independent variable indicates the length of time (measured in tenths of milliseconds) that elapsed from the presentation of the very first prime-target pair until the presentation of each of the subsequent pair until the very end of the experiment. Similar to RT, Time was recorded in tenths of milliseconds, but it was converted to milliseconds for the data analysis.

A Pearson product-moment correlation was run in order to examine the relationship between Time and RT. A weak yet significant negative correlation was found between the two variables, specifically, r(33,516) = − .08, $p < .001$. The correlation was also run with LANGUAGE as a split factor in order to verify whether the relationship manifested differently for both language groups. The weak negative correlation between Time and RT was maintained for both language groups, specifically, $r(16{,}464) = -.06$, $p < .001$ for bilinguals, and $r(17{,}052) = -.14$, $p < .001$ for functional monolinguals. While these results suggest that, irrespective of language group membership, participants were getting faster as the experiment progressed, the effect of Time on RT was stronger for functional monolingual participants.

With respect to accuracy, a point biserial correlation was run between Time and RAW_ACCURACY. The result of this correlation was not significant, $r(33{,}428) = .005$, $p = 401$. However, running it with LANGUAGE as a split factor yielded a significant result only for functional monolinguals, specifically, $r(17{,}010) = .03$, $p < .001$, suggesting, despite being weak, that functional monolinguals were improving (getting more accurate) as the experiment progressed.[8]

4.1.2.7. PRIME and TARGET

PRIME, as an independent variable, comprises each of the Portuguese primes in the experiment. Likewise, TARGET, comprises each of the English targets and pseudowords included in the experiment. Refer to Appendix F for the complete list of primes, targets, and pseudowords. Treating both PRIME and TARGET as independent variables allowed them to be entered into both mixed-effects models as random intercepts.

4.1.2.8. PAIR

Each of the prime-target pairs in the experiment were assigned a unique number code, which was then encoded by the variable PAIR. This allowed PAIR to be entered into the mixed-effects analyses as a random intercept. Refer to Appendix F for the unique number code assigned to each of the prime-target pairs.

---

[8] The point biserial correlation between Time and RAW_ACCURACY was not significant for bilinguals, $r(16{,}418) = .010$, $p = .179$.

4.1.2.9. PARTICIPANT

In order to account for the great deal of individual variation resulting from individual idiosyncrasies that cannot be easily (if not impossibly) controlled in an experiment like this, a subject variable must be included in a mixed-effects model analysis as a random intercept. To that end, as each participant in the experiment was assigned a pseudocode, the independent variable PARTICIPANT comprises the pseudocode assigned to each participant.

4.1.2.10. OTHER_LANGUAGES

All participants in the experiment had to indicate in the Language Background Questionnaire whether they had been exposed to (i.e., had received formal instruction) of another language. The variable OTHER_LANGUAGES encodes the participants' responses to this particular question as either YES or NO. As Figure 15 illustrates, the majority of the bilingual participants in the experiment (89.3%) answered NO to the question. This is in stark contrast with functional monolingual participants, whose majority (62.1%) answered YES to this question. This reflects the fact that most functional monolingual participants in the experiment were from the National Capital Region in Canada, an area where English speakers are exposed to French on a daily basis and often study the language in school as their second official language. Because of the great degree of orthographic, semantic and phonological overlap between French and Portuguese, both Romance

Languages, it was important to control for the participants' knowledge of another language to further understand its role in response latency. Thus, the variable OTHER_LANGUAGES was entered in both mixed-effects models as a control variable, which will be discussed in the next section.

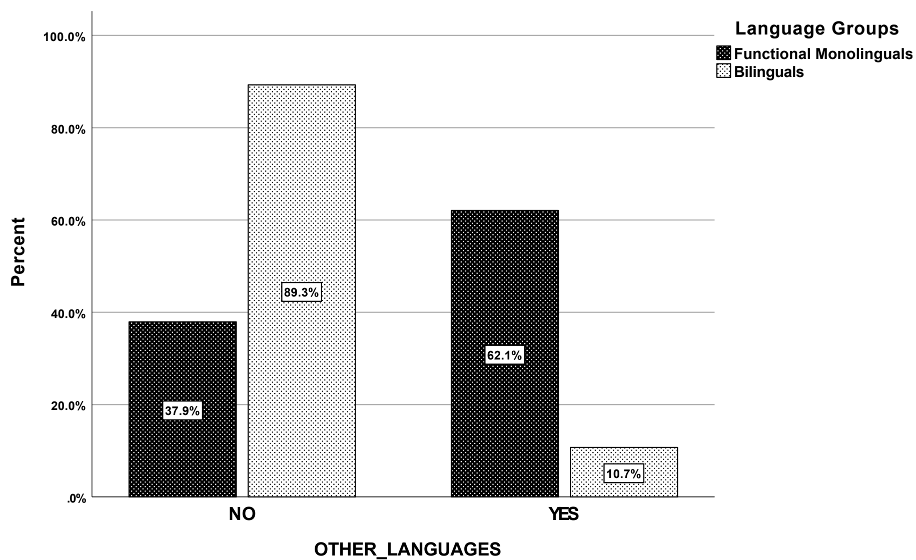


Figure 15: Differences between functional monolinguals and bilinguals with regards to whether or not they have the knowledge of another European language besides English and Portuguese.

4.1.3. Interim Summary of the Preliminary Analyses

Although the Preliminary Analysis section, as its name suggests, serves primarily as groundwork for the mixed-effects models that are presented in the ensuing section, an interim summary of the most import findings is provided below.

- With respect to overall RT, functional monolinguals were on average significantly faster (M = 730.79, SD = 276.08) than bilinguals (*M* = 808.60, *SD* = 348.29), $t(33{,}514) = -22.705$, $p < .001$.
- With respect to overall accuracy, functional monolinguals' percentage of correct answers (*M* = 95.50, *SD* = 3.08) was significantly higher than bilinguals' (*M* = 88.72, *SD* = 7.43), $t(33{,}426) = 109.54$, $p < .001$.
- A weak yet significant negative correlation between RT and RAW_ACCURACY, $r(33{,}428) = -.186$, $p < .001$, showed that on average as participants' accuracy increased, their reaction time decreased.
- With respect to English proficiency, based on SEMAC scoring, functional monolinguals were significantly more proficient in English than bilinguals, $t(33{,}514) = 119.97$, $p < .001$. Similarly, based on ERS scoring, functional monolinguals were more proficient in English than bilinguals, $t(33{,}514) = 116.88$, $p < .001$.
- An ANOVA with RT as the dependent variable and LANGUAGE and SEMANTIC as factors yielded a significant main effect of semantic similarity, $F(3, 33{,}508) = 875.88$, $p < .001$, a significant main effect of language group, $F(1, 33{,}508) = 333.88$, $p < .001$, as well as a significant interaction between semantic similarity and language group, $F(3, 33{,}508) = 10.06$, $p < .001$. Functional monolinguals were on average 73ms faster than bilinguals ($p < .001$). Participants from both language groups were on average 45ms faster to

respond to cognates and 194ms slower to respond to distractors in relation to controls (both *p*s $< .001$). Participants were also on average 29ms faster to respond to cognates than false friends ($p < .001$).

- Functional monolinguals' percentage of correct answers was higher than bilinguals' across all semantic categories. Distractor pairs yielded the lowest percentage of correct answers for both language groups while cognate pairs yielded the highest. These two findings indicate a possible main effect of LANGUAGE and SEMANTIC on RAW_ACCURACY.
- The degree of crosslinguistic orthographic overlap (NLD) between prime-target pairs classified as cognates and false friends is significantly higher than in pairs classified as controls and distractors, suggesting a possible main effect of NLD.
- A weak yet significant positive correlation between PHONETIC and RT was found, $r(33{,}516) = .02$, $p = .002$; on average participants reacted more slowly to more phonologically dissimilar pairs crosslinguistically. The correlation between PHONETIC and RAW_ACCURACY was not statistically significant ($p = .31$).
- The correlation between English_Freq and Portuguese_Freq was significant, $r(33{,}516) = .15$, $p < .001$. The correlation between RT and English_Freq was significant, $r(33{,}516) = -.11$, $p < .001$, and between RT and Portuguese_Freq, $r(33{,}516) = -.04$, $p < .001$. The correlation between both types of word

frequency and RAW_ACCURACY was significant for both English_Freq, $r(33{,}428) = .06$, $p < .001$, and Portuguese_Freq, $r(33{,}428) = .02$, $p < .001$.

- A significant correlation was found between RT and Time, $r(33{,}516) = -.08$, $p < .001$. The correlation was maintained for both language groups, specifically, $r(16{,}464) = -.06$, $p < .001$ for bilinguals, and $r(17{,}052) = -.14$, $p < .001$ for functional monolinguals. Despite the low *r*s, these correlations show that irrespective of their language group, participants' RT improved as the experiment progressed. With respect to RAW_ACCURACY, the correlation between this dependent variable with Time was not significant ($p = 401$). Using LANGUAGE as a split factor, the correlation between RAW_ACCURACY and Time was only significant for functional monolinguals, $r(17{,}010) = .03$, $p < .001$. Again, despite the weak *r*, this correlation shows that functional monolinguals were more accurate the experiment progressed.
- 89.3% of the bilingual participants indicated spoke no other European language besides English and Portuguese. In contrast, 62.1% of the functional monolingual participants indicated they spoke another European language besides English.

4.2. Mixed-Effects Models

The two dependent variables, viz., RT (response latency) and RAW_ACCURACY (response accuracy), were modeled separately as a function of all the independent variables introduced in the *Preliminary Analyses* section above. RT, as a continuous variable, was modeled using a linear mixed-effects model fitted with `lmer`. RAW_ACCURACY, as a dichotomous variable, was modeled using a logistic regression mixed-effects model with `glmer`. Both `lmer` and `glmer` are part of the `lme4` package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2019). Both mixed-effects models were fit in RStudio running remotely on a virtual Linux virtual instance using Amazon Elastic Compute Cloud (Amazon EC2)[9]. The results of each model are presented independently.

[9] Amazon Elastic Compute Cloud (Amazon EC2) is a subscription-based web service that is part of Amazon Web Services (AWS), a collection of remote computing services that together make up a cloud computing platform offered over the Internet by Amazon ("Amazon Elastic Compute Cloud (Amazon EC2)," 2018). Amazon EC2 provides secure, resizable compute capacity in the cloud, and offers a large number of instances, i.e., virtual computer configurations (Linux or Windows) that can be further customized to suit subscribers' computational and data processing needs. For the purposes of this study, Amazon EC2 was extremely useful during data analysis due to its computation power and processing speed, which significantly expedited model fitting in R Studio.

The random effects of PARTICIPANT, PAIR, PRIME, and TARGET were specified in both models. Their random intercepts were always included. Random slopes were included whenever possible; they were excluded if they correlated too highly ($r \geq .90$) with their random intercept or other random slopes. SEMANTIC, NLD, PHONETIC, English_Freq, Portuguese_Freq, and Time were entered as potential PARTICIPANT slopes.

A maximal mixed-effects structure was always sought for both models (Barr, Lev, Scheepers, & Tily, 2013). Iterations of the maximal model were successively simplified until the best fit was achieved. During the model simplification process, the most non-significant effect (highest *p*-value), starting with any of the specified interactions, was always removed first. This new iteration of the model was then refitted. If this iteration of the model successfully converged, subsequently, the anova function in R was called to compare the simplified (new iteration) model against its originator (previous iteration) model. The iteration that yielded the lowest Akaike Information Criterion (AIC) value (Akaike, 1998) continued to be further simplified in consecutive steps. The best model fit was deemed to have been reached when any further simplification to the model structure prevented it from successfully converging, produced a singularity fit, yielded a higher AIC value or was statistically the same its previous iteration (Chi-Square results provided in the anova).

In order to significantly minimize the occurrence of convergence errors when fitting both models, a few data cleaning techniques had to be applied. First, all continuous independent variables, viz., NLD, PHONETIC, English_Freq, Portuguese_Freq, and Time, were grand mean centered in order to reduce multicollinearity (Linck & Cunnings, 2015) and then standardized. Second, `nloptwrap2`, an optimizer function suggested by Bolker (2014) that calls for the *optimx* package (Nash, 2014) in R, was used to fit every iteration of both models. Finally, an inverse transformation (-1000/RT) was applied to the reaction time data in order to reduce its positive skew. Reaction time data, more often than not, particularly in language priming studies, usually turns out to be positively skewed. However, one of the major assumptions of several tests of statistical significance is that the dependent variable follows a normal (Gaussian) distribution. A positively skewed continuous dependent variable, as it is often the case with reaction time, violates that assumption. One of the suggested ways to mitigate this is to use invRT, which is applying an inverse transformation to the reaction time data (-1000/RT). Refer to Brysbaert & Stevens (2018) for further details on this technique.

Four random effects were always specified in both models: PARTICIPANT, PRIME, TARGET, and PAIR. Their random intercepts were always included. Random slopes with PARTICIPANT were included whenever possible. They were excluded if they correlated too highly ($r \geq .90$) with their random intercept or other random slopes.

The following fixed effects, as well as their interactions with one another, were specified in the `lmer` and `glmer` maximal models: LANGUAGE, SEMANTIC, NLD, PHONETIC. Several control variables were also included, specifically, LIST, English_Freq, Portuguese_Freq, target_length, prime_length, AGE, SEMAC (English proficiency), Time, and OTHER_LANGUAGES (whether or not participants had been exposed to or studied another European language besides English or Portuguese). As presented below, a number of fixed effects and control variables did not end up making to the best fit of both models, because they simply did not significantly improve model fit.

The coefficients of effect size (R^2 marginal and R^2 conditional) were computed for both models using the `r.squaredGLM` function included in the `MuMIn` package (Bartoń, 2019). R^2 marginal estimates the amount of variance accounted for by all the fixed effects included in the model (i.e., fixed effects, control variables, and interactions). R^2 conditional estimates the amount of variance accounted for by all the random effects and fixed effects included in the model (P. C. D. Johnson, 2014; Nakagawa & Schielzeth, 2013). Additionally, the *variance inflation factor* (VIF) and the VIF tolerances were computed for both models in order to assess multicollinearity between predictors (Field, Miles, & Field, 2012, p. 293). Both models yielded VIF values well below 10 and VIF tolerance values above .2. Based on these values, it can be safely concluded that there is no multicollinearity within the data (VIF values for both models are provided in the Appendix).

4.2.1 Mixed-Effects Models Results

```
LMER.BEST.FIT <- lmer(invRT ~ LANGUAGE*SEMANTIC + ZNLD_gmc*ZPHONETIC_gmc +
ZEnglish_Freq_gmc + Zlength_TARGET_raw_gmc + ZTime_ms_gmc + OTHER_LANGUAGES +
(1|PAIR) + (1|PRIME) + (1|TARGET) + (SEMANTIC|PARTICIPANT) +
(0+ZEnglish_Freq_gmc|PARTICIPANT) + (0+ZTime_ms_gmc|PARTICIPANT) +
(0+Zlength_TARGET_raw_gmc| PARTICIPANT), data = Dissertation_Dataset,
REML = FALSE, control=lmerControl(optimizer="nloptwrap2"))
```

4.2.1.1. Primary Fixed Effects

There was a significant main effect of LANGUAGE ($p < .01$). In comparison to functional monolingual participants (reference category), bilingual participants were on average 77ms slower. A significant main effect of SEMANTIC (semantic similarity) was also found for cognates ($p < .001$), false friends ($p < .01$) and distractors ($p < .001$). Specifically, irrespective of language group, in comparison to control pairs (reference category), cognate pairs yielded on average 22ms faster reaction times, false friend pairs yielded 14ms faster reaction times, and distractor pairs yielded 97ms slower reaction times. These findings are in-line with what was initially predicted; crosslinguistic semantic similarity or overlap on average had a significant effect on participants' RT.

4.2.1.2. Secondary Fixed Effects.

There was a significant main effect of English_Freq ($p < .001$). Specifically, participants' reaction time decreased on average six milliseconds as target frequency count increased. A significant main effect of target_length was also found ($p < .001$). Longer targets, those comprised of longer letter strings, slowed participants' reaction times 11ms on average. In addition, a main effect of Time was encountered ($p < .01$). Later trials yielded significantly faster reaction times than earlier ones, i.e., participants got on average 19ms faster as they progressed through the experiment. Finally, there was a trend towards a main effect of OTHER_LANGUAGES ($p = .07$). Specifically, participants with the knowledge of another language besides English and Portuguese were on average 45ms slower than participants who only spoke English or both English and Portuguese.

4.2.1.3. Interactions.

The LANGUAGE x SEMANTIC interaction was significant for cognate pairs ($p < .01$). Specifically, bilingual participants were on average 14ms faster on cognate pairs in comparison to functional monolinguals. In addition, the NLD (crosslinguistic orthographic overlap) x PHONETIC (crosslinguistic phonological overlap) interaction was also significant ($p < .01$). Specifically, prime-target pairs with a high degree of crosslinguistic orthographic overlap but with a low degree of crosslinguistic phonological overlap slowed down participants three milliseconds on average.

`LMER` model effect size.

The model's marginal effect size estimate was 16.14%. The primary and secondary fixed effects plus their interactions did not account for much of the observed variability in response latency. Nevertheless, the model's conditional effect size estimate was 61.43%. Adding the variance accounted for by PARTICIPANT, PAIR, PRIME and TARGET random effects explained over half of the observed variability.

Table 19 contains all of the `lmer` model results.[10]

[10] The raw RStudio output of the `lmer` model is provided in Appendix G.

Appendix G

Table 19. `lmer` Model Results

Fixed effects:	**invRT Estimate**	**RT ms Estimate**	**Effect in ms**	**Std. Error**	**df**	**t value**	***F***	**Pr(>|t|)**		***d***
(Intercept)	-1.713	584		0.064	70.91	-26.818	719.205	< 2E-16	***	
LANGUAGEBilinguals	0.199	660	77	0.076	64.02	2.601	6.765	0.011542	*	0.44
SEMANTICCOG	-0.068	562	-22	0.020	287.80	-3.451	11.909	0.000643	***	-0.15
SEMANTICDISTR	0.244	681	97	0.027	79.54	8.997	80.946	9.31E-14	***	0.54
SEMANTICFF	-0.041	570	-14	0.020	442.00	-2.044	4.178	0.041527	*	-0.09
ZNLD_gmc	-0.008	581	-3	0.006	1987.00	-1.295	1.677	0.195404		-0.02
ZPHONETIC_gmc	0.003	585	1	0.003	5485.00	1.267	1.605	0.205051		0.01
ZEnglish_Freq_gmc	-0.019	577	-6	0.005	461.70	-3.527	12.440	0.000462	***	-0.04
Zlength_TARGET_raw_gmc	0.030	594	11	0.006	366.70	5.121	26.225	4.92E-07	***	0.07
ZTime_ms_gmc	-0.057	565	-19	0.017	59.40	-3.389	11.485	0.00125	**	-0.13
OTHER_LANGUAGESYES	0.123	629	45	0.067	56.91	1.826	3.334	0.07305	.	0.27
LANGUAGEBilinguals:SEMANTICCOG	-0.042	570	-14	0.018	55.20	-2.32	5.382	0.023631	*	-0.09
LANGUAGEBilinguals:SEMANTICDISTR	-0.005	582	-2	0.035	56.56	-0.128	0.016	0.898777		-0.01
LANGUAGEBilinguals:SEMANTICFF	0.018	590	6	0.015	53.07	1.157	1.339	0.252305		0.04
ZNLD_gmc:ZPHONETIC_gmc	0.008	586	3	0.003	1671.00	2.299	5.285	0.021632	*	0.02

Random effects:

Groups	**Name**	**Variance**	**Std. Dev.**	**Corr.**		
PRIME	(Intercept)	0.0043107	0.06566			
PAIR	(Intercept)	0.0022357	0.04728			
TARGET	(Intercept)	0.0092168	0.096			
PARTICIPANT	Zlength_TARGET_raw_gmc	0.0004832	0.02198			
PARTICIPANT.1	ZTime_ms_gmc	0.0152537	0.12351			
PARTICIPANT.2	ZEnglish_Freq_gmc	0.0001005	0.01003			
PARTICIPANT.3	(Intercept)	0.0655086	0.25595			
	SEMANTICCOG	0.0031251	0.0559	0.46		
	SEMANTICDISTR	0.0170481	0.13057	-0.36	0.19	
	SEMANTICFF	0.001071	0.03273	0.54	0.63	0.22
Residual		0.0828356	0.28781			

4.2.2. Model 2 - Liner Mixed-Effects Model fitted with GLMER

```
GLMER.BEST.FIT <- glmer(RAW_ACCURACY ~ LANGUAGE*SEMANTIC + ZEnglish_Freq_gmc
+ ZTime_ms_gmc + Zlength_TARGET_raw_gmc + ZSEMAC_gmc + (1| PARTICIPANT) +
(1|PRIME) + (1|TARGET) + (1|PAIR), data = Dissertation_Dataset_LMER,
family=binomial, control=glmerControl(optimizer="nloptwrap2"))
```

4.2.2.1. Primary Fixed Effects

There was a significant main effect of LANGUAGE (p = .038). Functional monolinguals were overall more accurate than bilinguals. A main effect of SEMANTIC was also found (p < .001). Specifically, irrespective of their language group membership, participants were less accurate overall on distractor pairs.

4.2.2.2. Secondary Fixed Effects.

There was a significant main effect of English_Freq (p < .001). Specifically, irrespective of their language group membership, participants were more accurate in judging whether or not a target was a genuine English word when they were shown targets that occur more frequently in the language. A significant main effect of Time was also found (p < .001). Specifically, irrespective of their language group membership, participants' accuracy increased as they progressed through the task, i.e., they were more accurate at judging a target as a genuine English word when it occurred in later trials in the experiment. A main effect of target length

(length_TARGET) was also found ($p < .001$). Specifically, irrespective of language group membership, participants' accuracy increased as target length increased, i.e., they were more accurate in longer targets. Finally, a significant main effect of English proficiency (SEMAC) was found ($p < .001$). Specifically, irrespective of language group membership, participants who were on average more proficient in English, i.e., who had scored higher in the cloze proficiency test, were also more accurate in the experimental task.

4.2.2.3. Interactions.

The LANGUAGE x SEMANTIC interaction was significant ($p = .028$). Specifically, bilingual participants were more accurate than functional monolinguals controls in trials where the target were matched with a Portuguese prime cognate.

4.2.2.4. GLMER model effect size.

The model's marginal effect size estimate, which takes into account the response accuracy variability associated with all of the primary and secondary fixed effects plus their interactions, was 19.39%. The model's conditional effect size estimate was 56.25%. Adding the variance accounted for by PARTICIPANT, PAIR, PRIME and TARGET random effects explained over half of the observed variability.

Table 20 shows all of the `glmer` model results[11].

[11] The raw RStudio output of the `glmer` model is provided in Appendix H.

Table 20. `glmer` Model Results

	Estimates									
Fixed effects:	**Log Odds (Logit)**			**Odds Ratio**			**Std. Error**	**z value**	**Pr(>|z|)**	
	Log Odds (Logit)	**CI Lower Limit**	**CI Upper Limit**	**Odds Ratio**	**CI Lower Limit**	**CI Upper Limit**				
(Intercept)	4.877	-5.337	-4.418	131.248	0.005	0.012	0.234	20.817	< 2e-16	***
LANGUAGEBilinguals	-0.613	0.033	1.194	0.542	1.033	3.299	0.296	-2.073	0.038353	*
SEMANTICCOG	0.284	-0.740	0.172	1.329	0.477	1.187	0.233	1.222	0.221837	
SEMANTICDISTR	-1.372	0.997	1.748	0.253	2.710	5.744	0.192	-7.166	7.94E-13	***
SEMANTICFF	0.185	-0.713	0.343	1.203	0.490	1.410	0.269	0.686	0.492957	
ZEnglish_Freq_gmc	0.264	-0.431	-0.097	1.303	0.650	0.907	0.085	3.099	0.001942	**
ZTime_ms_gmc	0.174	-0.241	-0.108	1.191	0.786	0.898	0.034	5.136	2.80E-07	***
Zlength_TARGET_raw_gmc	0.220	-0.363	-0.077	1.246	0.695	0.926	0.073	3.017	0.002549	**
ZSEMAC_gmc	0.496	-0.753	-0.238	1.642	0.471	0.788	0.131	3.773	0.000162	***
LANGUAGEBilinguals:SEMANTICC	0.540	-1.022	-0.058	1.716	0.360	0.944	0.246	2.196	0.028062	*
LANGUAGEBilinguals:SEMANTICD	-0.136	-0.161	0.433	0.873	0.851	1.543	0.152	-0.900	0.368654	
LANGUAGEBilinguals:SEMANTICF	-0.348	-0.170	0.867	0.706	0.843	2.381	0.265	-1.317	0.18815	
Random effects:										
Groups	**Name**	**Variance**	**Std.Dev.**							
PRIME	(Intercept)	0.2766	0.5259							
PAIR	(Intercept)	0.3275	0.5723							
TARGET	(Intercept)	1.5102	1.2289							
PARTICIPANT	(Intercept)	0.6580	0.8112							

5. The Study: Discussion

Several well-known models of the mental lexicon leverage *spreading activation* to represent language schematically as a multilayered network of interconnected nodes, which is analogous to how cognitive psychologists posit human memory is organized (Anderson, 1983a, 1983b; Collins & Loftus, 1975; Collins & Quillian, 1969; Samani & Sharifian, 1997). In this framework, linguistic knowledge is depicted as network nodes; the associations between and among the countless types of linguistic knowledge are depicted as the connection links. Thus, when a network node is activated, it automatically spreads its activation, in iterative fashion, to other linked/associated nodes. Network nodes that are repeatedly activated together are represented closer to one another, with shorter connection links, in order to indicate their stronger linguistic relationships.

A hypothesis proposed by several well-known models of the bilingual mental lexicon, viz., RHM (J. F. Kroll & Stewart, 1994), BIA (Dijkstra & van Heuven, 1998; Van Heuven et al., 1998), BIA+ (Dijkstra & van Heuven, 2002) and Multilink (Dijkstra et al., 2018), is that bilingual language processing is non-selective. Specifically, during bilingual visual word recognition, lexical candidates from both L1 and L2 are simultaneously activated, which is a direct consequence of the bilingual mental lexicon being integrated, i.e., lexical representations from both languages are stored together.

The current study was conceived upon the principles set forth by spreading activation and concerned with Portuguese-English *cognates* and *false friends* with respect to their crosslinguistic form (orthographic and phonological) and meaning (semantic) overlap. From a theoretical perspective, its chief goal was to examine how lexical representations are organized in the Portuguese-English bilingual mental lexicon as well as to elucidate the lexical selection process. In addition, this study sought to quantify the specific effects that crosslinguistic form and meaning overlap, and any of their interactions, as well as word frequency, length, second language proficiency as control variables exert in this organization and during lexical selection. To that end, a lexical decision task with masked priming in the forward L1-to-L2 priming direction (Portuguese-English) was used. The experimental stimuli were comprised of Portuguese-English cognates and false friends. Portuguese primes used as experimental stimuli were also matched with English targets without form or meaning overlap crosslinguistically and served as experimental controls. Distractor items were comprised of Portuguese primes that had not been used as experimental or control stimuli and matched with pseudowords (nonwords) as targets. L1-Portuguese-L2-English bilinguals served as the experimental group and functional English monolinguals served as controls. Two mixed-effects analyses were carried out to analyze the data, and participants' response latency and accuracy served as dependent variables. The findings provide empirical support for many of the well-known models of the bilingual mental lexicon discussed in Chapter 2, in particular to the BIA+. Thus, the following discussion was structured primarily in regard to the

BIA+'s framework.

5.1. Findings

A main effect of LANGUAGE was observed for response latencies ($p < .05$). When other variables are not considered, bilingual participants' reaction times were significantly slower (77ms on average) than functional monolingual participants', indicating that lexical decision happens in a slower fashion overall for bilinguals. The main effect of LANGUAGE was also observed for response accuracy ($p < .05$). Bilingual participants' accuracy, i.e., deciding whether or not a target was an English word, was lower than functional monolingual participants'. These two results point to a language switching cost; when bilinguals are primed with an L1 word and the lexical decision is carried out in their L2, their reaction times are significantly slower relative to functional monolingual controls. Further, besides the cost of switching from their L1 to their L2, bilinguals' lexical decisions were less accurate compared to functional monolingual controls'. This decrease in lexical decision accuracy can be accounted for by Portuguese primes spreading their activation to other (non-target) English competitors. This reasoning is in line with the findings reported by von Studnitz & Green (1997), which also employed a masked priming lexical decision task.

A crosslinguistic priming effect was observed on participants' response latencies, but not on response accuracy. Both cognates (SEMANTICCOG) and false friends (SEMANTICFF) yielded on average faster reaction times in comparison to controls (both *p*s < .001). Cognates, however, had a significant 8ms speed-advantage over false friends. It follows then that crosslinguistic semantic overlap between a L1 prime and an L2 target exerts a facilitatory effect (22ms on average) on participants lexical decision times, potentially a result of the semantic representations exerting strong feedback activation on orthographic representations. It follows then that the greater the degree of semantic overlap between an L1 prime and L2 target, as it was the case with cognates, the faster participants' reaction times were on average. To further support this assertion, two other findings regarding crosslinguistic semantic overlap need also to be taken into consideration. Participants' reaction times in false friend trials (SEMANTICFF), i.e., word pairs in which the crosslinguistic form (orthographic and phonological) overlap between the Portuguese prime and English target was high but semantic overlap was low (or non-existent), despite being on average significantly faster than control trials ($p < .05$), were still less robust relative to cognate trials. Further, participants' reaction times on distractor trials (SEMANTICDISTR), i.e., pairs in which a bona fine Portuguese prime was matched to a pseudoword (nonword) target (no crosslinguistic form and meaning overlap whatsoever), were the slowest of all (97ms on average) in comparison to control trials ($p < .001$). Soares and Grosjean (1984), report similar findings with nonwords, with Portuguese-English bilinguals showing significantly slower reaction times to nonwords in comparison to genuine words from either language.

The aforementioned findings are, however, puzzling; bilinguals' and functional monolinguals' reaction times were significantly faster in cognate and false friend trials. This begs the question how functional monolingual participants could have been subject to crosslinguistic semantic priming, at least with respect to cognates, even when they expressed not having any knowledge of Portuguese during the intake assessment interview. Cognate and false friend trials have a common characteristic: a high degree of crosslinguistic form (orthographic and phonological) overlap compared to unrelated controls. It is thus possible that upon presentation of the masked Portuguese prime, its lexical and sublexical orthographic (and phonological) representations may have been sufficiently able to activate a number of lexical and sublexical orthographic and phonological representations in both the target and non-target language, which, in turn, facilitated participants' lexical decisions, even in the absence of shared semantic representations to exert feedback activation. Further, since the effect was also significant in false friend trials irrespective of language group, it is safe to assume that crosslinguistic orthographic priming was in fact the culprit.

The interaction between language group and crosslinguistic semantic overlap (LANGUAGEBilinguals:SEMANTICCOG) was significant, i.e., a facilitatory crosslinguistic priming effect (often reported in the literature as *cognate facilitation effect*) was observed on bilinguals' response latencies and response accuracy (both *p*s $< .05$). Bilinguals were on average 14ms faster to respond to an English target that was semantically, orthographically, and phonologically related to a matched Portuguese prime (cognate) than functional monolingual controls. In addition, bilingual participants were more accurate at deciding whether a target as a bona fide English word when it was matched to a semantically, orthographically and phonologically related Portuguese prime relative to an unrelated Portuguese control prime, which corroborates the results reported in Alves-Soares (2013). Because bilingual participants were not only significantly faster to react but also significantly more accurate to respond to English targets in cognate trials compared with functional monolingual participants, it is safe to conclude that both semantic and crosslinguistic form overlap facilitate bilinguals' lexical decisions.

These findings align well with the results of several noteworthy studies on the bilingual mental lexicon (Altarriba & Basnight-Brown, 2007; Caramazza & Brones, 1979; Cristoffanini et al., 1986; Dijkstra et al., 1999; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Duñabeitia, Perea, & Carreiras, 2010; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Forster et al., 2003; Gerard, Linda, & Scarborough, 1989; Guasch et al., 2013; Kim & Davis, 2003; Lemhöfer & Dijkstra, 2004; Nakayama et al., 2012; Peeters, Dijkstra, & Grainger, 2013; Sanchez-Casas, Davis, & Garcia-Albea, 1992; Schwartz et al., 2007). Taking into account the framework proposed by the BIA+ and its updated version named the Multilink (both discussed in Chapter 2), they also have major implications for the inner workings as well as the organization of the Portuguese-English bilingual mental lexicon. First, these findings show that crosslinguistic priming effects are automatic and not under direct control of the bilingual. Second, a certain level of crosslinguistic integration exists in the bilingual mental lexicon. Specifically, presentation of a prime, even when it is masked, in one language automatically activates a number of candidates in the target language proportional to the degree of crosslinguistic orthographic overlap between the prime and target. Third, when the degree of crosslinguistic form and semantic overlap between an L1 prime and an L2 target is high, i.e., when they are cognates, their shared semantic representations exert strong feedback activation on their orthographic representations; thus, the cognate facilitation effect is a direct result of cognates receiving increased activation from their shared semantics, which accelerate bilinguals' responses for this group of words. In addition, in the absence of crosslinguistic semantic overlap between an L1 prime and an L2 target, facilitatory crosslinguistic priming effects can still occur due to crosslinguistic form overlap. Finally, these findings support for the language non-selective access hypothesis, i.e., lexical candidates from bilinguals' both languages are simultaneously activated upon presentation of word in either language during the first stages of visual word recognition. With respect to the current study, presentation of a masked Portuguese prime facilitated bilinguals' recognition of a cognate English target.

Despite the fact that crosslinguistic orthographic (ZNLD_gmc) and phonological overlap (ZPHONETIC_gmc) individually yielded no significant effect on participants' response times and accuracy, their interaction

(ZNLD_gmc:ZPHONETIC_gmc) was significant ($p < .05$) on participants' response latencies. This interaction, however, did not enter the second mixed-effects model (logistic regression `glmer` model), as it did not contribute significantly to participants' response accuracy.

In order to better understand how crosslinguistic form overlap affected participants' response latencies, some facts regarding stimuli selection need to be recollected (see chapter 3, Methodology for further detail on stimuli selection). The Normalized Levenshtein Distance (Schepens et al., 2013) was selected as an objective metric of crosslinguistic orthographic overlap. It yields values that range between 0, when orthographic overlap between a prime and its matched target is nonexistent, and 1, when the prime and target are orthographically identical. Dynamic time warping was selected as an objective metric of crosslinguistic phonological overlap, and Phonological Corpus Tools (Hall et al., 2016) was the software package used to implement the DTW algorithm. These two objective metrics allowed this study to more elegantly intuit how these two variables interact in the Portuguese-English bilingual mental lexicon. In addition, examining this interaction in a mixed-effects analysis with this language pair had not been yet attempted.

The significant interaction between crosslinguistic orthographic and phonological overlap suggests that a certain level of integration exists between both the L1 and the L2 and that orthographic representations spread their activation to associated phonological representations. Specifically, during the course of visual word recognition, upon presentation of an L1 prime, orthographic representations of the prime become activated and spread their activation in a feed-forward fashion to orthographic representations from both languages based on their degree of orthographic overlap with the prime. Likewise, phonological representations from both languages are also activated subsequently based on their degree of phonological overlap with the prime. In addition, this significant interaction between crosslinguistic orthographic and phonological overlap may indicate the BIA+'s "temporal delay assumption" (Dijkstra & van Heuven, 2002, p. 183) deserves merit. The model proposes that there are time course differences in the contribution of the two codes in lexical decision, with lexical and sublexical orthographic code being activated first and L2 lexical and sublexical phonological code being activated slightly later. As such, crosslinguistically phonological overlap appears to modulate activation spread, particularly in the forward L1-to-L2 direction. It follows then the significant 3ms inhibitory effect observed with this interaction may be a consequence of the temporal delay. Alternatively, it is possible that this 3ms delay is a result of the inconsistencies in mapping Portuguese orthographic code to English phonological code (and vice-versa). To that end, a certain grapheme (or its combinations) in Portuguese may be consistently pronounced the same way in the language, whereas in English, the pronunciation of the grapheme is much less consistent. These differences in crosslinguistic phonological overlap relative to orthographic overlap may induce lateral inhibition, since the non-overlapping phonological code would also activate lexical representations that are less phonologically similar to both prime and target, which would reduce the magnitude/robustness of the facilitatory crosslinguistic form overlap effect. Although both these explanations can be accommodated within the BIA+ framework, it is important to point out that it is rather difficult (perhaps nearly impossible) to precisely disentangle crosslinguistic phonological effects from orthographic effects in languages such as Portuguese and English (as well as French, Spanish and Italian), which despite sharing the same roman script and having a great number of words with 100% orthographic overlap, almost never have words with 100% phonological overlap.

One of the control variables that was of particular interest in this study was word frequency. While Portuguese (prime/L1) word frequency (ZPortuguese_Freq_gmc) yielded no significant main effect on both response latencies and accuracy, probably a consequence of the experiment having been conducted in the forward L1-to-L2 priming direction (Portuguese-English), English (target/L2) word frequency (ZEnglish_Freq_gmc) did have an effect on both of dependent variables. Participants, irrespective of language group membership, were on average 6ms faster ($p < .001$) to select targets that occur more frequently in English, i.e., have a higher frequency count per million words in the language. In addition, irrespective of their language group membership, i.e., whether they were classified bilingual or functional monolingual, participants were more accurate at recognizing high frequency count targets as genuine words of the English language, compared to low frequency targets ($p < .01$).

These two findings provide further insight of the inner workings of the Portuguese-English bilingual mental lexicon. To that end, L2 word frequency facilitates lexical decision in terms of speed and accuracy, but this facilitatory effect is not substantial (only a 6ms facilitation in response times, $d = .03$). Both groups of participants were significantly faster and more accurate on average to respond to English targets that occur more frequently in the language. This finding is in line with the BIA+ proposal. The model purports to say that word frequency affects how activation spreads in the Word Identification System, which, in turn, affects how the Task Schema selects a specific target. Specifically, frequency of use modulates how activation spreads from the input string to linked representations in the Word Identification System by raising (or even lowering) the resting-level activation of stored representations. In this fashion, representations linked to frequently used words have sufficiently higher resting-level activation and, consequently, spread their activation to other linked representations much faster, which, during the course of visual word recognition, facilitates lexical decision. Likewise, frequently used L2 targets have a higher baseline (resting) level of activation, i.e., they are more salient, thus, less activation is required from an L1 prime to effect activation of L2 target candidates and the spread of activation between prime and target occurs at a significantly faster rate.

The BIA+ accounts for how representations from one language can have sufficiently higher resting-level activation relative to representations from the other language. In respect to the current study, because English was the target language of the lexical decision task and was also the primary language of communication for all of the participants, it would be safe to assume that English representations had ample time to establish themselves in the mental lexical, i.e., increase their resting-level activation, and exert strong feedback activation on other competitor English representations, which facilitated the Task Schema's verification (checking) whether a specific target belonged to English, even when competing Portuguese representations might have been available for the bilingual participants. In this framework, as bilingual and functional monolingual participants were using English as their primary language of communication, study and work on a daily basis, and the lexical decision task was carried out in English, it is possible that "recency of use of the target language" not only accelerated lexical selection in English but also exerted top-down inhibition that made lexical selection in the language significantly more accurate. Had the experimental task been conducted in the backwards priming direction (English prime – Portuguese target), it is conceivable that L1 word frequency effects on both response latency and accuracy would have been detected, but only for bilingual participants. In order to more precisely quantify the effects of recency of use, it would be important for future studies to compare bilinguals living in their L2 with bilinguals living their L1 in future studies.

Since the effect of L2 word frequency was shown not to be substantial, instead of adding word frequency as an isolated control variable in a mixed-effects model, as it was done in the current study, future studies should assemble an experimental stimuli list in which the frequency of the readings of cognates and false friends is carefully controlled for relative to controls and pseudowords. Additionally, it would be important verify whether switching the direction of the lexical decision task, viz., Portuguese prime-English target vs. English prime-Portuguese-target vs. Portuguese prime-Portuguese target vs. English-prime vs. English target, yields substantive differences.

Target length (Zlength_TARGET_raw_gmc) is another control variable that had a significant effect on participants' response latency. Participants, irrespective of language group membership, processed longer targets, i.e., were comprised of more letters, on average 11ms slower ($p < .001$). Besides potentially having been influenced by the priming direction of the experimental task, this finding suggests a temporal delay/cost (inhibitory effect) associated with the processing of longer words, potentially due to greater memory (cognitive) load requirements. In this way, upon the presentation of an L1 prime, a number of lexical orthographic representations are simultaneously activated dependent on their degree of orthographic of overlap with prime during the first states of visual word recognition. As prime-target pairs in the experimental task were matched for length, and both languages of the experiment share the same alphabet, it is possible that the number of lexical competitors (or neighbors) activated by the L1 prime is a function of the total number of letters that comprises the prime. Specifically, longer Portuguese primes activated a greater number of lexical competitors in English (the target language) due to an increased number of acceptable letter permutations crosslinguistically, slowing down lexical selection by 11ms on average.

Target word length also had a significant effect on response accuracy, irrespective of language group membership. Participants were more accurate at recognizing a target as a genuine English word when they were longer in length ($p < .01$). This shows that despite the cost associated with the simultaneous activation of English competitors, which slowed participants down 11ms on average, lexical selection still occurred in a more accurate fashion. This finding provides support for the existence a *task schema*, an idea first purported in the BIA+. According to Dijkstra & van Heuven (2002), the *Task Schema*, which is in constant communication with the *Word Identification System*, is responsible for ensuring proper "safe execution" of the task at hand by verifying the membership of possible lexical candidates within a specific language. Because the effect of target word length affected both bilinguals and functional monolinguals, it would be safe to assume a task schema had to properly verify the membership of all the active lexical competitors in the target language, in this case English, in order to accurately narrow down a specific target. As word length affected the number of possible competitors simultaneously active, a greater number of lexical competitors allowed the task schema to more accurately select an English target, thereby yielding the significant effect of response accuracy observed.

The variable ZTime_ms_gmc, which encoded the amount of time participants, irrespective of language group membership, spent on the experimental task, yielded a significant effect on both response latency ($p < .01$) as well as response accuracy ($p < .001$). Bilinguals' and functional monolinguals' lexical decisions were on average faster (19ms) and more accurate on later trials compared to earlier ones.

Two scenarios may explain these findings. (a) A facilitatory effect of practice: as participants got more familiar with the lexical decision task, their reaction times and accuracy in later trials significantly improved, probably also due to the rather large total number of trials that comprised the experiment. In fact, in the extant literature the effect of practice has been observed in crosslinguistic lexical decision tasks with respect to response latency and accuracy when data was analyzed using mixed effects models (Keuleers, Diependaele, & Brysbaert, 2010). Nevertheless, despite the highly significant ($p < .001$) effect of practice, in practical terms it was described as rather small, (40ms difference in response latency and 2% difference in response accuracy) and deemed to be a result of the large number of trials. In the current study, the effect of practice may have been partly an upshot of the large number of trials as well. In addition, the observed effect was also rather small (19ms difference in response latency by the end of the task). (b) A facilitatory recency effect: the linguistic representations of earlier prime-target pairs may have retained a proportion of their activation and facilitated the activation of linguistic representations of prime-target pairs in later trials, thereby expediting the lexical selection process. As Portuguese primes in the experimental condition were matched with English targets that were unrelated in both form and meaning, it is reasonable to assume that despite pseudorandomization, Portuguese primes from earlier trials could have retained a proportion of their activation and facilitated lexical selection of targets in later trials. In addition, it is plausible that this recency effect was further accentuated (additive effect) by the fact that all participants were using English as their primary language of communication and that the experimental task carried out in the Portuguese-English direction, with English being the target language. In other words, living in an English-speaking environment may have given English targets an edge, allowing participants to process them faster and more accurately.

English proficiency had no impact on response latency, and, in fact, it did not enter the `lmer` model. Nevertheless, English proficiency (ZSEMAC_gmc) yielded a main effect that was highly significant for response accuracy ($p < .001$), irrespective of language group membership, and thus, entered the `glmer` model. Specifically, participants who obtained a higher score on the cloze proficiency test were on average significantly more accurate at judging a target as a genuine word of the English language.

Several studies have reported robust translation priming effects (asymmetry) in the forward L1-L2 priming direction using masked priming technique in unbalanced bilinguals (Altarriba & Basnight-Brown, 2007; Dimitropoulou, Duñabeitia, & Carreiras, 2011c; Finkbeiner, Forster, Nicol, & Nakamura, 2004; Jiang, 1999; Nakayama, Ida, & Lupker, 2016; Nakayama, Sears, Hino, & Lupker, 2013; Wang, 2013; Wen & van Heuven, 2017; Xia & Andrews, 2015). The robust translation asymmetries are suggested to occur due to an imbalance between bilinguals' L1 and L2. Specifically, L1 lexical representations have a higher resting-level activation compared to L2 lexical representations, thus activation spreads more quickly and easily in the forward L1-to-L2 masked priming direction than the other way around. In addition, it has been reported that in order to obtain similar effects in the backwards L2-to-L1 direction, unbalanced bilinguals require that the L2 prime be presented for longer, greater than the standard Stimuli Onset Asynchrony (SOA) of 50ms (Wen & van Heuven, 2017).

Nevertheless, it has also been reported in the literature that the translation priming effect is less robust or completely disappears for balanced bilinguals (Duñabeitia, Dimitropoulou, et al., 2010; Duñabeitia, Perea, et al., 2010; Perea et al., 2008). Additionally, the BIA+ model (Dijkstra & van Heuven, 2002), proposes that as L2 proficiency increases, the resting-level activation of L2 lexical representation also increases. As a result, translation priming asymmetry effects for balanced bilinguals as well as highly proficient bilinguals would not be significant or, if significant, their effect size (d) would be small (Wen & van Heuven, 2017).

The aforementioned findings indicate that the bilinguals who participated in the current study may have been more proficient in English than the cloze proficiency test was able to detect. In addition, it is likely that using their L2 (English) as their primary language of communication in Canada on a daily basis affected the balance between the L1 and the L2, placing English in a more equal position to their L1 (Portuguese). These two factors explain why English proficiency had no effect on response latency for both bilinguals and functional monolinguals. Regarding the significant effect of English proficiency on response accuracy, from a theoretical standpoint, this finding is compatible with the BIA+ framework. Specifically, daily exposure to their L2 not only helped bilinguals to become more proficient in English but also allowed them to develop more stable lexical representations in the language (higher resting-level activation), in turn, improving lexical decision accuracy. In addition, it is likely that more stable L2 representations exert stronger top-down inhibition suppressing the activation of competitor lexical representations in the non-target language leading to better overall accuracy (Dijkstra & van Heuven, 2002, p. 189).

Finally, the control variable OTHER_LANGUAGES yielded a main effect that approached statistical significance (p = .08) with respect to response latencies, thus deserves further examination. The variable encoded whether or not participants spoke another European language besides English and Portuguese, which for the purposes of this study was important to account for. Many models of the bilingual mental lexicon (in particular the BIA and its updated version BIA+) propose that the bilingual lexicon is integrated, that lexical access occurs in a non-selective fashion, and that during the first stages of visual word recognition a number of orthographic candidates from bilinguals' both languages become activated depending on their orthographic similarity with the input string (Van Heuven et al., 1998). This (priming) effect is hypothesized to be proportional to the degree of crosslinguistic form overlap (Cristoffanini et al., 1986). Therefore, participants that speak languages with a high degree of crosslinguistic form overlap, such as French and Portuguese, may manifest slower reaction times on a crosslinguistic lexical decision task. To that end, this (almost significant) main effect of OTHER_LANGUAGES is potentially a result of the high degree of crosslinguistic orthographic overlap between Portuguese and French, both Romance Languages, and between both of these languages and English. In the current study, all of the selected primes were in Portuguese. Upon presentation of the Portuguese prime, Portuguese, English (and possibly French) candidates were simultaneously activated for the bilingual participants. The Portuguese prime activated English and potentially French candidates for functional monolingual participants. Since the experiment took place at the University of Ottawa, an institution where students are exposed to both English and French regularly, in Canada's National Capital Region, an area where residents are also exposed to both English and French regularly, it is likely that lexical decision was slowed down by their knowledge of French. Despite not being the target language, French is historically related to Portuguese; consequently, both languages share a great degree of orthographic code. Thus, having knowledge of French, or having been exposed to it regularly, may have exerted an inhibitory effect on participants' reaction times. Furthermore, it is also possible that a few functional monolingual participants may have learned other Romance Languages, such as Spanish and Italian, whose words are be more similar orthographically to Portuguese than French. Ultimately, this result indicates that more research is required in order to cognize how crosslinguistic orthographic overlap between a prime and a non-target, yet historically similar, third language can affect bilinguals' lexical decision.

6. Conclusion and Suggestions for Further Research

In sum, the aforementioned findings provide strong empirical support for the BIA+. First, the significant condition effects encountered suggest some level of integration in the bilingual mental lexicon. Specifically, presentation of a Portuguese prime facilitated participants' reaction times to an English target when crosslinguistic form (orthographic and phonological) overlap between them was high (cognates and false friends). Because the effect was significant irrespective of language group and crosslinguistic form overlap in cognates and false friends is high, it follows that crosslinguistic form overlap was likely the responsible factor. Second, crosslinguistic semantic overlap plays a leading role in the organization of the bilingual mental lexicon. When the Portuguese prime and the English target shared the same conceptual representation in permanent memory (cognates), bilingual participants' reaction times were significantly faster than when the Portuguese prime and the English target did not share the same conceptual representation (false friends). This effect suggests that cognates may have a special representation in the bilingual mental lexicon. Specifically, besides sharing a single semantic representation crosslinguistically, cognates may share a single morphemic representation comprised of common lexical and sublexical orthographic plus phonemic representations in both bilingual's languages. Besides providing strong evidence in favor integration, the findings of this study also support the BIA+'s depiction of the bilingual mental lexicon as being hierarchically organized into two cognitive systems, a Word

Identification System, which sits atop the hierarchy, and a Task Decision/Schema System. The aforementioned crosslinguistic priming effects observed with bilingual participants can be accounted for as being the result of these two complex, yet highly interactive, cognitive processes. To that end, during the course of visual word recognition, lexical and sublexical orthographic representations that comprised a masked Portuguese prime automatically activated a number of similar (linked) lexical and sublexical orthographic representations simultaneously in both Portuguese and English, which, subsequently activated a number of similar lexical and sublexical phonological representations in both languages as well. As English was bilingual participants' primary language of communication, English word frequency, recency of use as well as proficiency level in the language had increased the resting-level activation of words in the language relative to Portuguese words, which allowed activation to spread to other similar English representations much faster. Semantic representations further facilitated or inhibited activation spread to similar lexical and sublexical orthographic and phonological representations in both languages, which competed against one another. Recognition of the target ultimately occurred when the Task Decision/Schema System, after continuous weighing of activation spread in the Word Identification System and taking into account the lexical decision task's instructions, determined that the target had met recognition threshold.

Despite all of the evidence gathered in support of the BIA+, the findings above warrant future study in other to address the limitations encountered. First, the pool of bilingual participants was not uniform. A sizeable number of bilingual participants was comprised of international undergraduate students from Brazil who had come to the University of Ottawa to participate in a one-year undergraduate exchange program in their respective fields of study. Only a few of the bilingual participants reported to be pursuing, or already had, post-graduate level of education. In addition, while all bilingual participants indicated during the intake assessment interview that English had become their primary language since arriving in Canada, most of them had learned the language after starting high school in Brazil, which is usually the norm in the country. Moreover, many of the bilingual participants admitted to having studied French; some even mentioned they had attended *L'Alliance Française* in Brazil for many years while they were also learning English. Adding to this, many reported they were living in Gatineau, Québec, and had francophone roommates, i.e., were exposed to French on a daily basis. Although one of the original goals of this study was to quantify the effect of language dominance, age of English acquisition and manner of acquisition on both dependent variables, aforementioned circumstances made it impossible for this to be carried out.

Similarly, despite the fact that our functional monolingual participants were native speakers of English and did not speak or know Portuguese, living in Ottawa, an officially bilingual city, and working at University of Ottawa, the largest bilingual university in Canada, meant that this group of participants was still exposed to

French on a daily basis. Some of the functional monolingual participants also reported having studied Spanish and Italian. Since Portuguese, Spanish, Italian and French are Romance Languages and share a great deal of morphological and phonological similarities, it is possible these very facts skewed the results by making functional monolingual participants more similar to bilingual participants. Therefore, it would be critical for future studies to be stricter in selecting participants for the control group, i.e., separating native speakers of English who speak only English from native speakers of English who speak other romance languages, and not conflate both into one single group.

Second, few studies have shed light on the dynamics of the Task Decision Schema and the Word Identification System, in particular with respect to how the types pseudowords (or nonwords) can sway how activation threshold is crossed for recognition of a word to take place during a lexical decision task (Grainger & Jacobs, 1996; Lima & Huntsman, 1997; Lupker, Brown, & Colombo, 1997; Taylor & Lupker, 2001). To that end, after decision parameters are set during practice trials for an English-only lexical decision task, the Task Decision Schema can be adapted or fine-tuned over the course of the experiment, which can significantly skew target recognition. Consequently, participants' reaction times to pseudoword targets may be facilitated, or participants may start automatically selecting a NO response for pseudoword targets, if, for instance, all of the pseudowords are orthographically or even phonologically unpronounceable in the target language. Despite the fact the results of this study demonstrated that both group of participants were slowest on average to respond to pseudoword targets (relative to unrelated English controls), in order to minimize potential stimuli composition effects, it is vital for future studies to continue to exercise restraint and select only pseudowords that are orthographically and phonologically legal in English for distractor pairs. In addition, it is vital that pseudowords match Portuguese primes in length and that their orthographic and phonological overlap is low, i.e., they must not be interlingual homographs or homophones.

Finally, although using objective metrics to assess the effects of crosslinguistic overlap on the organization of the Portuguese-English bilingual mental lexicon demonstrated to be a steppingstone that yielded meaningful results, future studies should also factor in subjective metrics in mixed-effects model analyses. For instance, it would be productive to establish whether there are significant differences in model effect size between objective and subjective metrics, or even whether it would be possible to create a composite metric, combining subjective and objective crosslinguistic overlap metrics together.

In conclusion, this dissertation found empirical support for the BIA+ model, in particular with respect to the language non-selective hypothesis, which suggest that linguistic representations from bilinguals' both languages were activated during the course of visual word recognition. Additionally, even though crosslinguistic orthographic and phonological overlap alone did not facilitate bilinguals' recognition of English targets, their interaction did. Specifically, bilingual participants, compared with functional monolingual participants, were on average significantly faster to respond and more accurate to identify English targets that were matched with orthographically and phonologically similar Portuguese primes. Additionally, this effect was further facilitated when the Portuguese prime and the English target shared the same semantic representation, i.e., when they were cognates. These findings suggest that cognates may share a single morphemic representation for in the Portuguese-English bilingual mental lexicon. Moreover, it was found that both bilingual and functional monolingual participants were sensitive to frequency of use and word length in English only.

Taking into account the framework proposed by the BIA+, crosslinguistic overlap facilitates how bilinguals (relative to functional monolingual controls) process cognates. The significant crosslinguistic semantic priming effect observed with this group of participants and cognates showed that crosslinguistic overlap indeed facilitates bilinguals' lexical decisions. Specifically, because bilingual participants were not only significantly faster to react but also significantly more accurate to respond to cognates compared with unrelated controls than functional monolingual participants did, it is safe to conclude that lexical decisions are facilitated by both semantic and crosslinguistic orthographic overlap. In addition, it appears that semantic similarity in fact plays a major role in how activation spreads crosslinguistically, since in its absence, as it was the case with false friends, no crosslinguistic priming effect was observed. The BIA+ is able to accommodate the significant effects of target word length on both response times and accuracy, as well as the effect of time spent in the experiment.

This dissertation contributes to the advancement of the linguistic and psycholinguistic fields in three major ways. First, despite the great deal of crosslinguistic similarity or semantic, orthographic and phonological code overlap between Portuguese and English, little research has been directed at expounding how Portuguese-English bilinguals process cognates and false friends in the two languages. In fact, the role that both crosslinguistic similarity and English proficiency play in the organization of the Portuguese-English bilingual mental lexicon has been largely overlooked. To that end, this dissertation serves as a steppingstone in addressing these issues in a more quantitative fashion and also provides undeniable evidence in support of the BIA and BIA+. Second, from a methodological standpoint, this dissertation contributes to a core issue in bilingualism by demonstrating that crosslinguistic similarity can be effectively measured using objective metrics of similarity such as dynamic time warping (DTW) and the normalized Levenshtein distance (NLD); thus, the hurdles and limitations many second language researchers encounter for relying on subjective metrics of crosslinguistic similarity can now be successfully overcome. Third, although ANOVAs and t-tests have been the primary statistical analysis (means-based parametric) tools utilized in many second language studies to analyze bilingual data collected in LDTs, liner mixed-effects models were chosen instead for this dissertation. Besides being more robust and more precise at estimating the individual effects of crosslinguistic similarity on both dependent variables (reaction time and accuracy), mixed-effects models yielded a more realistic representation of the processes involved in visual word recognition. Finally, at the applied level, the fields of translation and language pedagogy could undoubtedly benefit from the results presented in this dissertation by incorporating cognates in teacher training and in the preparation of pedagogical materials.

7. References

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4612-1694-0_15

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic- and ..., *39*(1), 1–18.

Alves-Soares, L. (2013). *Investigating the English-Portuguese Bilingual Mental Lexicon: Effects of Orthography and Phonology*. University of Ottawa, Ottawa, ON.

Amazon Elastic Compute Cloud (Amazon EC2). (2018). Amazon Web Services, Inc. Retrieved from https://aws.amazon.com/ec2/

Anderson, J. R. (1983a). A Spreading Activation Theory of Memory. *Journal of Verbal Learning and Behavior*, *22*, 261–265. https://doi.org/10.1016/B978-1-4832-1446-7.50016-9

Anderson, J. R. (1983b). *The Architecture of Cognition*. Cambridge, MA, USA: Harvard University Press.

Anderson, J. R. (2015). *Cognitive psychology and its implications* (8th ed.). New York, NY: Worth Publishers.

Audacity®: Free Audio Editor and Recorder. (2019). Audacity Team. Retrieved from https://audacityteam.org/

Barr, D. J., Lev, R., Scheepers, C., & Tily, H. J. (2013). Keep it maximal Appendix.

Journal of Memory and Language, *68*(3), 1–5. https://doi.org/10.1016/j.jml.2012.11.001.Random

Bartoń, K. (2019). MuMIn: Multi-Model Inference. https://doi.org/10.1111/j.1420-9101.2010.02210.x

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bolker, B. (2014). Convergence error for development version of lme4. Retrieved October 30, 2017, from https://stackoverflow.com/questions/21344555/convergence-error-for-development-version-of-lme4

Bradlow, A., Clopper, C., Smiljanic, R., & Walter, M. A. (2010). A perceptual phonetic similarity space for languages: Evidence from five native language listener groups. *Speech Communication*, *52*(11–12), 930–942. https://doi.org/10.1016/j.specom.2010.06.003

Brown, J. D. (1980). Relative Merits of Four Methods for Scoring Cloze Tests. *The Modern Language Journal*, *64*(3), 311–317. https://doi.org/10.1111/j.1540-4781.1980.tb05198.x

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1(1)*(9), 1–20. https://doi.org/10.5334/joc.10

Brysbaert, Marc. (2003). Bilingual visual word recognition: Evidence from masked phonological priming. In *Masked Priming: The State of the Art* (pp. 178–188). https://doi.org/10.4324/9780203502846

Brysbaert, Marc, & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism*, *13*(3), 359–371. https://doi.org/10.1017/S1366728909990344

Brysbaert, Marc, Van Dyck, G., & Van De Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/0096-1523.25.1.137

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, *13*(4), 212–214. https://doi.org/10.3758/BF03335062

Chen, H. C., & Leung, Y. S. (1989). Patterns of Lexical Processing in a Nonnative Language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/0278-7393.15.2.316

Chomsky, N., & Halle, M. (1992). *The Sound Pattern of English. PsycCRITIQUES*. Cambridge, MA.: MIT Press.

Clements, G., & Hume, E. V. (1995). The internal organization of speech sounds. In J.

Goldsmith (Ed.), *The handbook of phonological theory* (pp. 245–306). Cambridge, MA.

Clements, G. N. (1985). The geometry of phonological features. *Phonology*, *2*(1985), 225. https://doi.org/10.1017/S0952675700000440

Collins, A. M., & Loftus, E. F. (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Behavior*, *8*, 240–247. https://doi.org/10.1016/S0022-5371(69)80069-1

Colombo, L. (1986). Activation and Inhibition With Orthographically Similar Words. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/0096-1523.12.2.226

Comesaña, M., Sánchez-Casas, R., Soares, A. P., Pinheiro, A. P., Rauber, A., Frade, S., & Fraga, I. (2012). The interplay of phonology and orthography in visual cognate word recognition: An ERP study. *Neuroscience Letters*, *529*(1), 75–79. https://doi.org/10.1016/j.neulet.2012.09.010

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The Cognate Facilitation Effect: Implications for Models of Lexical Access. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/0278-7393.26.5.1283

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism*, *9*(2), 137–151. https://doi.org/10.1017/S1366728906002495

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*, 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of spanish-english cognates. *The Quarterly Journal of Experimental Psychology Section A*, *38*(3), 367–393. https://doi.org/10.1080/14640748608401604

da Luz Fontes, A. B. A., & Schwartz, A. I. (2010). On a different plane: Cross-language effects on the conceptual representations of within-language homonyms. *Language and Cognitive Processes*, *25*(4), 508–532. https://doi.org/10.1080/01690960903285797

Davies, M. (2017a). Corpus of Contemporary American English. Retrieved from https://corpus.byu.edu/coca/

Davies, M. (2017b). O corpus do Português. Retrieved from http://www.corpusdoportugues.org/x.asp

de Groot, A. M. B. (1992). Determinants of Word Translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/0278-7393.18.5.1001

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*(4), 496–518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301. https://doi.org/10.1016/j.jml.2009.12.003

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. *Localist Connect. Approaches to Hum. Cogn.*, 189–225.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/s1366728902003012

Dijkstra, T., Wah, BUYTENHUIJS, F., VAN HALEM, N., AL-JIBOURI, Z., DE KORTE, M., & REKKÉ, S. (2018). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011a). Masked translation priming effects with low proficient bilinguals. *Memory and Cognition*, *39*(2), 260–275. https://doi.org/10.3758/s13421-010-0004-9

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011b). Phonology by itself: Masked phonological priming effects with and without orthographic overlap. *Journal of Cognitive Psychology*, *23*(2), 185–203. https://doi.org/10.1080/20445911.2011.477811

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011c). Two words, one meaning: Evidence of automatic co-activation of translation equivalents. *Frontiers in Psychology*, *2*(AUG), 1–20. https://doi.org/10.3389/fpsyg.2011.00188

Duñabeitia, J. A., Dimitropoulou, M., Uribe-Etxebarria, O., Laka, I., & Carreiras, M. (2010). Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. *Brain Research*, *1359*, 142–154. https://doi.org/10.1016/j.brainres.2010.08.066

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology*, *57*(2), 98–107. https://doi.org/10.1027/1618-3169/a000013

Dunabeitia, J. A., Perea, M., Carreiras, M., Duñabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology*, *57*(2), 98–107. https://doi.org/10.1027/1618-3169/a000013

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual Word Recognition by Bilinguals in a Sentence Context: Evidence for Nonselective Lexical Access. *Journal of Experimental Psychology: Learning Memory and*

Cognition, *33*(4), 663–679. https://doi.org/10.1037/0278-7393.33.4.663

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society*. https://doi.org/10.3758/BF03334157

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London, England: SAGE Publications Ltd.

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, *51*(1), 1–22. https://doi.org/10.1016/j.jml.2004.01.004

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/0278-7393.10.4.680

Forster, K. I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In *Masked Priming: The State of the Art* (pp. 2–20). Psychology Press. https://doi.org/10.4324/9780203502846

Friesen, D. C., & Jared, D. (2012). Cross-language phonological activation of meaning: Evidence from category verification. *Bilingualism*, *15*(1), 145–156. https://doi.org/10.1017/S1366728910000489

Gerard, Linda, & Scarborough, D. (1989). Language-Specific Lexical Access of Homographs by Bilinguals - Tags: HOMONYMS SEMANTICS. *Journal of Experimental Psychology. Learning; Memory & Cognition*, *15(2)*, 305–315.

Retrieved from http://connection.ebscohost.com/c/articles/14387219/language-specific-lexical-access-homographs-by-bilinguals

Gollan, T. H., Forster, K. I., & Frost, R. (1997). crucial for Translation Priming With Different Scripts : Masked Priming With providing and Noncognates in Bilinguals the effect of. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1122–1139. Retrieved from http://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=1997-06939-004&S=L&D=pdh&EbscoContent=dGJyMMvl7ESeqa44v%2BbwOLCmr1Cep7NSsKu4TLCWxWXS&ContentCustomer=dGJyMPGot1CzrrZLuePfgeyx44Dt6fIA

Grainger, J., & Frenck-Mestre, C. (2002). Masked Priming by Translation Equivalents in Proficient Bilinguals. *Language and Cognitive Processes*, *13*(6), 601–623. https://doi.org/10.1080/016909698386393

Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Linguistics and Language Compass*. https://doi.org/10.1111/j.1749-818X.2008.00121.x

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-Out Model Word recognition is an integral component of language pro- cessing in general. It raises a fundamental question about the. *Psychological Review Carr & Pollatsek*, *103*(3), 518–565.

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychological Science*.

https://doi.org/10.1111/j.1467-9280.2006.01821.x

Grainger, J., Spinelli, E., Farioli, F., Diependaele, K., & Ferrand, L. (2003). Masked Repetition and Phonological Priming within and across Modalities. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(6), 1256–1269. https://doi.org/10.1037/0278-7393.29.6.1256

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, *45*(3), 765–771. https://doi.org/10.3758/s13428-012-0296-8

Hall, K. C., Allen, B., Fry, M., Mackie, S., & McAuliffe, M. (2016). Phonological CorpusTools: A free, open-source tool for phonological analysis. *14th Conference for Laboratory Phonology*. Retrieved from http://corpustools.readthedocs.io/en/latest/index.html

Hutchison, K. A. (2007). Attentional Control and the Relatedness Proportion Effect in Semantic Priming. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/0278-7393.33.4.645

IBM SPSS Statistics 25. (2018). Armonk, NY: IBM Corp.

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition*, *5*(03). https://doi.org/10.1017/S1366728902003024

Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition*, *2*(1), 59–75.

https://doi.org/10.1017/s1366728999000152

Johnson, K. (2011). *Acoustic and Auditory Phonetics* (3rd ed.). Malden, MA: Wiley-Blackwell.

Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R2GLMMto random slopes models. *Methods in Ecology and Evolution*, *5*(9), 944–946. https://doi.org/10.1111/2041-210X.12225

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633. https://doi.org/10.3758/BRM.42.3.627

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology*, *1*(NOV), 1–15. https://doi.org/10.3389/fpsyg.2010.00174

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language*, *49*(4), 484–499. https://doi.org/10.1016/S0749-596X(03)00093-7

Kondrak, G., & Sherif, T. (2006). Evaluation of several phonetic similarity algorithms on the task of cognate identification. *Proceedings of the Workshop on Linguistic Distances*, 43–50. https://doi.org/10.3115/1641976.1641983

Kroll, J., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In *Practical Aspects of Memory*.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language*, *33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., & Tokowicz, N. (2001). The development of conceptual representation for words in a second language. *One Mind, Two Lang. Biling. Lang. Process.*, 49–71.

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory and Cognition*, *32*(4), 533–550. https://doi.org/10.3758/BF03195845

Levelt, W. J. M. (1993). *Speaking: From Intention to Articulation (ACL-MIT Series in Natural Language Processing)*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (2001). Spoken word production: a theory of lexical access. In *Proceedings of the National Academy of Sciences* (Vol. 98 (23), pp. 13464–13471). Nijmegen, The Netherlands. https://doi.org/10.1073/pnas.231459498

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75. https://doi.org/10.1017/S0140525X99001776

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*. https://doi.org/citeulike-article-id:311174

Lima, S. D., & Huntsman, L. A. (1997). Sequential dependencies in the lexical decision task. *Psychological Research*, *60*(4), 264–269. https://doi.org/10.1007/BF00419412

Linck, J. A., & Cunnings, I. (2015). The Utility and Application of Mixed-Effects Models in Second Language Research. *Language Learning*, *65*(S1), 185–207. https://doi.org/10.1111/lang.12117

Litz, D. R., & Smith, A. K. (2006). Semantically Acceptable Scoring Procedures (SEMAC) Versus Exact Replacement Scoring Methods (ERS) For ' Cloze ' Tests : A Case Study. *Asian EFL Journal Quaterly*, *8*(1), 1–28. Retrieved from https://www.asian-efl-journal.com/1133/quarterly-journal/semantically-acceptable-scoring-procedures-semac-versus-exact-replacement-scoring-methods-ers-for-cloze-tests-a-case-study/#squelch-taas-tab-content-0-0

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic Control in a Naming Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(3), 570–590.

Martin, R. C., & Jensen, C. R. (1988). Phonological priming in the lexical decision task: A failure to replicate. *Memory & Cognition*. https://doi.org/10.3758/BF03197052

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*. https://doi.org/10.1037/0033-295X.88.5.375

Mielke, J. (2012). A phonetically based metric of sound similarity. *Lingua*, *122*(2), 145–

163. https://doi.org/10.1016/j.lingua.2011.04.006

Montrul, S. A. (n.d.). *Incomplete Acquisition in Bilingualism: Re-Examining the Age Factor*. Philadelphia, PA, USA: John Benjamins Publishing Co.

Motley, M. T., & Camden, C. T. (1985). Nonlinguistic influences on lexical selection: Evidence from double entendres. *Communications Monographs*, *52*(2), 124–135. https://doi.org/10.1080/03637758509376100

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nakayama, M., Ida, K., & Lupker, S. J. (2016). Cross-script L2-L1 noncognate translation priming in lexical decision depends on L2 proficiency: Evidence from Japanese-English bilinguals. *Bilingualism*, *19*(5), 1001–1022. https://doi.org/10.1017/S1366728915000462

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2012). Cross-script phonological priming for Japanese-English bilinguals: Evidence for integrated phonological representations. *Language and Cognitive Processes*, *27*(10), 1563–1583. https://doi.org/10.1080/01690965.2011.606669

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2013). Masked translation priming with Japanese-English bilinguals: Interactions between cognate status, target frequency and L2 proficiency. *European Journal of Cognitive Psychology*, *25*(8), 949–981. https://doi.org/10.1080/20445911.2013.839560

Nash, J. C. (2014). On Best Practice Optimization Methods in {R}. *Journal of Statistical Software*, *60*(2), 1–14. Retrieved from http://www.jstatsoft.org/v60/i02/

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The Influence of Native-Language Phonology on Lexical Access: Exemplar-Based Versus Abstract Lexical Entries. *Psychological Science*, *12*(6), 445–449. https://doi.org/10.1111/1467-9280.00383

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language*, *68*(4), 315–332. https://doi.org/10.1016/j.jml.2012.12.003

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language*, *58*(4), 916–930. https://doi.org/10.1016/j.jml.2008.01.003

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*. https://doi.org/10.1016/0749-596X(91)90017-E

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning Memory and Cognition*, *24*(1), 101–118. https://doi.org/10.1037/0278-7393.24.1.101

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone Effects in Lexical

Decision. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*(1), 139–156. https://doi.org/10.1037//0278-7393.27.1.139

Pollatsek, A., Perea, M., & Carreiras, M. (2005). Does conal prime CANAL more than cinal? Masked phonological priming effects in Spanish with the lexical decision task. *Memory and Cognition*, *33*(3), 557–565. https://doi.org/10.3758/BF03193071

Potter, M. C., So, K. F., Eckardt, B. Von, & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*. https://doi.org/10.1016/S0022-5371(84)90489-4

Presentation®: Precise, Powerful Stimulus Delivery. (2017). Berkeley, CA: Neurobehavioral Systems, Inc. Retrieved from https://www.neurobs.com

R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria. https://doi.org/10.1007/978-3-540-74686-7

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*. https://doi.org/10.1080/01690960050119689

RStudio Team. (2018). RStudio: Integrated Development Environment for R. Boston, MA. Retrieved from http://www.rstudio.com/

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*.

https://doi.org/10.1037/0033-295X.89.1.60

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. https://doi.org/10.1017/S1366728913000643

Samani, R., & Sharifian, F. (1997). Cross-language hierarchical spreading of activation. In F. Sharifian (Ed.), *Conference on Language, Cognition, and Interpretation* (pp. 11–23). Isfahan: IAU Press. Retrieved from http://cogprints.org/1663/00/act2.htm

Sanchez-Casas, R. M., Davis, C. W., & Garcia-Albea, J. E. (1992). Bilingual Lexical Processing: Exploring the Cognate/ Non-cognate Distinction. *European Journal of Cognitive Psychology*, *4*(4), 293–310. https://doi.org/10.1080/09541449208406189

Schepens, J., Dijkstra, T., Grootjen, F., & van Heuven, W. J. B. (2013). Cross-Language Distributions of High Frequency and Phonetically Similar Cognates. *PLoS ONE*, *8*(5). https://doi.org/10.1371/journal.pone.0063006

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*(2), 197–212. https://doi.org/10.1016/j.jml.2006.03.004

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive*

Processes, *22*(1), 106–129. https://doi.org/10.1080/01690960500463920

Shinde, R. B., & Pawar, V. P. (2014). Dynamic time Warping using MATLAB & PRAAT. *International Journal of Scientific & Engineering Research*, *5*(5), 37–42.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition*, *12*(4), 380–386. https://doi.org/10.3758/BF03198298

Sunderman, G., & Schwartz, A. I. (2008). Using Cognates to Investigate Cross-Language Competition in Second Language Processing. *TESOL Quarterly*, *42*(3), 527–536. https://doi.org/10.1002/j.1545-7249.2008.tb00145.x

Taylor, T. E., & Lupker, S. J. (2001). Sequential Effects in Naming: A Time-Criterion Account. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*(I), 117–138.

The World's Top 10 Largest Economies. (n.d.). Retrieved October 1, 2019, from https://www.focus-economics.com/blog/the-largest-economies-in-the-world

Tzur, B., & Frost, R. (2007). *SOA does not Reveal the Absolute Time Course of Cognitive Processing in Fast Priming Experiments*. *J Mem Lang* (Vol. 56).

van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*(3), 193–211. https://doi.org/10.1017/s1366728998000352

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/BF03196335

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*(3), 181–198. https://doi.org/10.3758/BF03197716

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/0096-1523.28.3.616

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition*, *35*(5), 938–952.

von Studnitz, R. E., & Green, D. W. (1997). Lexical decision and language switching. *International Journal of Bilingualism*, *1*(1), 3–24. https://doi.org/10.1177/136700699700100102

Wang, X. (2013). Language dominance in translation priming: Evidence from balanced and unbalanced Chinese-English bilinguals. *Quarterly Journal of Experimental Psychology*, *66*(4), 727–743. https://doi.org/10.1080/17470218.2012.716072

Weber, R. (1970). A Linguistic Analysis of First-Grade Reading Errors. *Reading*

Research Quarterly, *5*(3), 427–451. Retrieved from http://www.jstor.org/stable/747079

Wen, Y., & van Heuven, W. J. B. (2017). Non-cognate translation priming in masked priming lexical decision experiments: A meta-analysis. *Psychonomic Bulletin and Review*, *24*(3), 879–886. https://doi.org/10.3758/s13423-016-1151-1

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Retrieved from http://ggplot2.org

Xia, V., & Andrews, S. (2015). Masked translation priming asymmetry in Chinese-English bilinguals: Making sense of the Sense Model. *Quarterly Journal of Experimental Psychology*, *68*(2), 294–325. https://doi.org/10.1080/17470218.2014.944195

Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., & Grainger, J. (2000). Visual and phonological codes in letter and word recognition: Evidence from incremental priming. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*. https://doi.org/10.1080/713755906

8. Appendix

Appendix A. Research Ethics Approval Forms

File Number: 09-12-13 **Date (mm/dd/yyyy):** 02/10/2015

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Ethics Approval Notice

Social Sciences and Humanities REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Juana	Munoz-Liceras	Arts / Modern Languages and Literatures	Supervisor
Marc	Brunelle	Arts / Linguistics	Co-Supervisor
Leonardo	Alves-Soares	Arts / Linguistics	Student Researcher

File Number: 09-12-13

Type of Project: PhD Thesis

Title: Investigating the Effects Cross-Language Activation during Cross-Modal Language Processing

Renewal Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
01/22/2015	01/21/2016	Ia

(Ia: Approval, Ib : Approval for initial stage only)

Special Conditions / Comments:
N/A

1

550, rue Cumberland, pièce 154 Ottawa (Ontario) K1N 6N5 Canada
550 Cumberland Street, room 154 Ottawa, Ontario K1N 6N5 Canada
(613) 562-5387 • Téléc./Fax (613) 562-5338
www.recherche.uottawa.ca/deontologie/ www.research.uottawa.ca/ethics/

File Number: 09-12-13 **Date (mm/dd/yyyy):** 02/10/2015

Université d'Ottawa University of Ottawa

Bureau d'éthique et d'intégrité de la recherche Office of Research Ethics and Integrity

This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement (2010) and other applicable laws and regulations in Ontario, has examined and approved the ethics application for the above named research project. Ethics approval is valid for the period indicated above and subject to the conditions listed in the section entitled "Special Conditions / Comments".

During the course of the project, the protocol may not be modified without prior written approval from the REB except when necessary to remove participants from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the project (e.g., change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, including consent and recruitment documentation, should be submitted to the Ethics Office for approval using the "Modification to research project" form available at: http://research.uottawa.ca/ethics/submissions-and-reviews.

Please submit an annual report to the Ethics Office four weeks before the above-referenced expiry date to request a renewal of this ethics approval. To close the file, a final report must be submitted. These documents can be found at: http://research.uottawa.ca/ethics/submissions-and-reviews.

If you have any questions, please do not hesitate to contact the Ethics Office at extension 5387 or by e-mail at: ethics@uOttawa.ca.

Signature:

Mélanie Rioux
Ethics Coordinator
For Catherine Paquet, Director of the Office of Research Ethics and Integrity

550, rue Cumberland, pièce 154 550 Cumberland Street, room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5387 • Téléc./Fax (613) 562-5338
www.recherche.uottawa.ca/deontologie/ www.research.uottawa.ca/ethics/

File Number: 09-12-13 **Date (mm/dd/yyyy):** 04/24/2017

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Ethics Approval Notice

Social Sciences and Humanities REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Juana	Munoz-Liceras	Arts / Modern Languages and Literatures	Supervisor
Marc	Brunelle	Arts / Linguistics	Co-Supervisor
Leonardo	Alves-Soares	Arts / Linguistics	Student Researcher

File Number: 09-12-13

Type of Project: PhD Thesis

Title: Investigating the Effects Cross-Language Activation during Cross-Modal Language Processing

Renewal Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
01/22/2016	01/21/2017	Renewal

Special Conditions / Comments:
N/A

550, rue Cumberland, pièce 154 550 Cumberland Street, room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5387 • Téléc./Fax (613) 562-5338
www.recherche.uottawa.ca/deontologie/ www.research.uottawa.ca/ethics/

File Number: 09-12-13 **Date (mm/dd/yyyy):** 04/24/2017

Université d'Ottawa University of Ottawa
Bureau d'éthique et d'intégrité de la recherche Office of Research Ethics and Integrity

Ethics Approval Notice

Social Sciences and Humanities REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

First Name	Last Name	Affiliation	Role
Juana	Munoz-Liceras	Arts / Modern Languages and Literatures	Supervisor
Marc	Brunelle	Arts / Linguistics	Co-Supervisor
Leonardo	Alves-Soares	Arts / Linguistics	Student Researcher

File Number: 09-12-13

Type of Project: PhD Thesis

Title: Investigating the Effects Cross-Language Activation during Cross-Modal Language Processing

Renewal Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
01/22/2017	01/21/2018	Renewal

Special Conditions / Comments:
N/A

550, rue Cumberland, pièce 154 550 Cumberland Street, room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5387 • Téléc./Fax (613) 562-5338
www.recherche.uottawa.ca/deontologie/ www.research.uottawa.ca/ethics/

File Number: 09-12-13 **Date (mm/dd/yyyy):** 04/24/2017

Université d'Ottawa University of Ottawa
Bureau d'éthique et d'intégrité de la recherche Office of Research Ethics and Integrity

This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement (2010) and other applicable laws and regulations in Ontario, has examined and approved the ethics application for the above named research project. Ethics approval is valid for the period indicated above and subject to the conditions listed in the section entitled "Special Conditions / Comments".

During the course of the project, the protocol may not be modified without prior written approval from the REB except when necessary to remove participants from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the project (e.g., change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, including consent and recruitment documentation, should be submitted to the Ethics Office for approval using the "Modification to research project" form available at: https://research.uottawa.ca/ethics/forms.

Please submit an annual report to the Ethics Office four weeks before the above-referenced expiry date to request a renewal of this ethics approval. To close the file, a final report must be submitted. These documents can be found at: https://research.uottawa.ca/ethics/forms.

If you have any questions, please do not hesitate to contact the Ethics Office at extension 5387 or by e-mail at: ethics@uOttawa.ca.

Signature:

Mélanie Rioux
Ethics Coordinator
For Catherine Paquet, Director of the Office of Research Ethics and Integrity

550, rue Cumberland, pièce 154 550 Cumberland Street, room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5387 • Téléc./Fax (613) 562-5338
www.recherche.uottawa.ca/deontologie/ www.research.uottawa.ca/ethics/

File Number: 09-12-13

Date (mm/dd/yyyy): 04/24/2017

Université d'Ottawa University of Ottawa

Bureau d'éthique et d'intégrité de la recherche Office of Research Ethics and Integrity

This is to confirm that the University of Ottawa Research Ethics Board identified above, which operates in accordance with the Tri-Council Policy Statement (2010) and other applicable laws and regulations in Ontario, has examined and approved the ethics application for the above named research project. Ethics approval is valid for the period indicated above and subject to the conditions listed in the section entitled "Special Conditions / Comments".

During the course of the project, the protocol may not be modified without prior written approval from the REB except when necessary to remove participants from immediate endangerment or when the modification(s) pertain to only administrative or logistical components of the project (e.g., change of telephone number). Investigators must also promptly alert the REB of any changes which increase the risk to participant(s), any changes which considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project and safety of the participant(s). Modifications to the project, including consent and recruitment documentation, should be submitted to the Ethics Office for approval using the "Modification to research project" form available at: https://research.uottawa.ca/ethics/forms.

Please submit an annual report to the Ethics Office four weeks before the above-referenced expiry date to request a renewal of this ethics approval. To close the file, a final report must be submitted. These documents can be found at: https://research.uottawa.ca/ethics/forms.

If you have any questions, please do not hesitate to contact the Ethics Office at extension 5387 or by e-mail at: ethics@uOttawa.ca.

Signature:

Mélanie Rioux
Ethics Coordinator
For Catherine Paquet, Director of the Office of Research Ethics and Integrity

550, rue Cumberland, pièce 154 550 Cumberland Street, room 154
Ottawa (Ontario) K1N 6N5 Canada Ottawa, Ontario K1N 6N5 Canada
(613) 562-5387 • Téléc./Fax (613) 562-5338
www.recherche.uottawa.ca/deontologie/ www.research.uottawa.ca/ethics/

Appendix B. Participant Recruitment Form

INVITATION TO PARTICIPATE IN A LANGUAGE STUDY

Target Participants

- Portuguese - English bilinguals (all proficiency levels)
- English - Portuguese bilinguals (all proficiency levels)
- English monolinguals

Description

You are being invited to participate in a Cross-Modal Language Processing study, which is being conducted by Leonardo Alves-Soares, a Ph.D. student in the Department of Linguistics at the University of Ottawa and supervised by professor Juana Muñoz-Liceras. The purpose of the study is to further understand cross-language activation during the processing of words with varying degrees of orthographic and phonetic overlap across the two languages of a bilingual. The study will take a approximately 90 minutes of your time, and it will take place either in the Sound Patterns Laboratory, Simard Hall, room 333 (3rd floor), or in the Second Language Acquisition Research Laboratory, Arts Hall, room 253, which are located at the University of Ottawa main campus.

Your participation in the study is completely voluntary.

You will not receive compensation for your participation in the study. However, if you wish, we can provide you with a certificate of participation, which you can include with your CV.

Contact

If you are interested in participating in the study, or if you have any questions before deciding whether or not to participate, please contact the experimenter, Leonardo Alves-Soares

Appendix C: Informed Consent Forms

uOttawa

Université d'Ottawa
Faculté des arts
Linguistique

University of Ottawa
Faculty of Arts
Linguistics

613-562-5286
613-562-5141
70 Laurier E.
Ottawa ON K1N 6N5 Canada
www.uOttawa.ca

INFORMED CONSENT FORM

Experimenter: Leonardo Alves-Soares

Contact Information: Department of Linguistics
University of Ottawa
Hamelin Hall, 70 Laurier Ave. E., Room 415
Ottawa, ON K1N 6N5

Supervisors: Professor Juana Muñoz-Liceras
Department of Modern Languages and
Department of Linguistics
University of Ottawa
Arts Hall, 70 Laurier Ave. East, room 217
Ottawa, ON K1N 6N5

Professor Marc Brunelle
Department of Linguistics
University of Ottawa
Arts Hall, 70 Laurier Avenue East, room 429

Invitation to Participate
You are invited to participate in a study titled *Investigating the Portuguese-English Bilingual Mental Lexicon: Effects of Orthography and Phonology.*

Purpose of the Study
The purpose of the study is to investigate cross-language activation during the processing of words across the two languages of a bilingual.

Participation
The following is the sequence of events that will take place during the study.

(1) You will fill out a language background questionnaire.

(2) You will take an English proficiency test (cloze).

(3) You will be asked to sit at a computer workstation for the experimental task. Your accuracy and reaction time will be recorded for analysis.

uOttawa

Université d'Ottawa
Faculté des arts
Linguistique

University of Ottawa
Faculty of Arts
Linguistics

Risks
This study does not involve risk or deception.

Confidentiality and Anonymity
Your name and contact information will remain strictly confidential. Your data will be identified by code only. All data linked to you will match the code previously assigned to you. The data will not contain your name or any other biographical information.

Conservation of Data
The data will be kept forever. All (digital) data collected for the study will be stored in the investigator's main computer, which is password-protected.

The original consent forms as well as their copies will be stored in secured filing cabinet in the Sound Patterns Laboratory (SMD 333). No one, besides the experimenter himself and/or his supervisors, will have access to them.

Further Use of Data
All data collected in the study could be reanalyzed by other researchers to verify the validity of the results or compare them with the results of other studies. In such a case, your data will be transmitted anonymously, and all your personal information will be kept confidential. Please indicate your preference by **circling** one of the two options below.

You agree that your data be used in further analyses

You disagree that your data be used in further analyses.

Voluntary Participation
You are under no obligation to participate in the study. If you choose to participate, you may withdraw at any time. If you choose to withdraw, all data gathered up to the time of withdrawal will be discarded. No negative consequences will result from either
your refusal to participate or withdraw.

613-562-5286
613-562-5141
70 Laurier E.
Ottawa ON K1N 6N5 Canada
www.uOttawa.ca

uOttawa

Université d'Ottawa
Faculté des arts
Linguistique

University of Ottawa
Faculty of Arts
Linguistics

Acceptance
By signing below, you acknowledge that you understand the terms, parameters and conditions of this research study conducted by Leonardo Alves-Soares, Ph.D. student (the "Experimenter") in the Department of Linguistics at the University of Ottawa. This study is under supervision of Professor Juana Muñoz-Liceras and Professor Marc Brunelle (the "Supervisors").

If you have any questions about the study, you may contact either the Experimenter or the Supervisor using the contact information provided in the beginning of this form. Alternatively, you may also contact the Protocol Officer for Ethics in Research, University of Ottawa, Tabaret Hall, 550 Cumberland Street, Room 154, Ottawa, Ontario, K1N 6N5, 613-562-5387, or at ethics@uottawa.ca.

There are two copies of this consent form, one of which is yours to keep.

Participant

Print Name Signature Date

Experimenter

Print Name Signature Date

613-562-5286
613-562-5141
70 Laurier E.
Ottawa ON K1N 6N5 Canada
www.uOttawa.ca

uOttawa

Université d'Ottawa
Faculté des arts
Linguistique

University of Ottawa
Faculty of Arts
Linguistics

FORMULÁRIO DE CONSENTIMENTO

Investigador Principal: Leonardo Alves-Soares
(estudante de doutorado)
Departamento de Linguística
Universidade de Ottawa

Contato: Departamento de Linguística
Universidade de Ottawa
Pavilhão Artes, Avenida Laurier Leste nº 70, 4º andar

Ottawa, ON K1N 6N5

Professores Supervisores:
Professora Juana Muñoz-Liceras
Departmento de Línguas Modernas e
Departmento de Linguística
Universidade de Ottawa
Pavilhão Artes
Avenida Laurier Leste nº 70, sala 217
Ottawa, ON K1N 6N5

Professor Marc Brunelle
Departmento de Linguística
Universidade de Ottawa
Pavilhão Artes
Avenida Laurier Leste nº 70, sala 429
Ottawa, ON K1N 6N5

Convite para Participação
Você está sendo convidado para participar no estudo intitulado *Investigating the Effects Cross-Language Activation during Cross-Modal Language Processing.*

Propósito do Estudo
O propósito do estudo é investigar a ativação simultânea das duas línguas de um bilíngue durante processamento de palavras cognatas.

Participação
Durante o estudo, a seguinte cadeia de eventos irá acontecer.

(1) Você preencherá um questionário com questões sobre o seu background linguístico.

613-562-5286
613-562-5141
70 Laurier E.
Ottawa ON K1N 6N5 Canada
www.uOttawa.ca

uOttawa

Université d'Ottawa
Faculté des arts
Linguistique

University of Ottawa
Faculty of Arts
Linguistics

613-562-5286
613-562-5141
70 Laurier E.
Ottawa ON K1N 6N5 Canada
www.uOttawa.ca

2) Participante bilíngues farão dois testes de proficiência, um na língua nativa e o outro na segunda língua.

(3) Assim que você terminar o teste de proficiência, você irá sentar-se em frente a uma workstation e fará uma tarefa sobre decisões léxicas. A sua acuidade e tempos de respostas serão medidos.

(4) Após o experimento, participantes bilíngues serão dados uma lista de palavras em sua segunda língua e eles terão que traduzir essas palavras para a sua língua nativa.

Riscos
Este estudo não envolve riscos ou decepção.

Confidencialidade e Anonimato
O seu nome e informações pessoais permanecerão estritamente confidenciais. Seus dados serão somente identificáveis por um código primeiramente associado a você, e todos dados coletados durante este estudo serão correlacionados com esse código. Os dados coletados não conterão o seu nome ou sequer outra informação bibliográfica sua.

Conservação de Dados
Os dados coletados durante este estudo jamais serão destruídos. Todos dados (digitais) coletados serão arquivados no computador pessoal do pesquisador principal, que é protegido por senha, e uma cópia será armazenada num pen drive criptografado, que será guardado no escritório dos supervisores no campus da universidade.

Os formulários de consentimento e suas cópias serão arquivados num arquivo no laboratório. Ninguém, a não ser o pesquisador principal e seus professores supervisores, terão acesso a eles.

Uso dos Dados Posteriormente
Todos os dados coletados neste estudo poderão ser reanalisados por outros pesquisadores a fim de que possam verificar a validade dos resultados deste estudo ou compará-los com resultados de outros estudos. Em tal caso, seus dados serão transmitidos anonimamente, e todas suas informações pessoais permanecerão confidenciais. Por favor, indique sua preferência circulando uma das duas opções abaixo.

Você concorda que seus dados poderão ser utilizados em análises futuras.

Você discorda que seus dados poderão ser utilizados em análises futuras.

uOttawa

Université d'Ottawa
Faculté des arts
Linguistique

University of Ottawa
Faculty of Arts
Linguistics

Participação Voluntária
Você não está sob nenhuma obrigação sequer de participar neste estudo. Se você decidir em participar, você pode rescindir sua participação em qualquer momento. Se você decidir rescindir sua participação, todos os seus dados coletados serão descartados. Nenhuma consequência negativa sequer ocorrerá como resultado de você recusar-se de participar neste estudo ou rescindir sua participação

Aceita
Assinando abaixo, você concorda que você entende os termos, parâmetros e condições deste estudo conduzido por Leonardo Alves-Soares, estudante de doutorado (o "Investigador Principal") no Departamento de Linguística na Universidade de Ottawa, estudo supervisionado pelos professores Juana Muñoz-Liceras e Marc Brunelle (os "Professores Supervisores").

Se você tiver qualquer pergunta sobre este estudo, favor contatar o investigador principal ou os professores supervisores usando as informações de contato providenciadas no começo deste documento. Alternativamente, você pode contatar o Oficial de Protocolo de Ética em Pesquisas pessoalmente no pavilhão Tabaret, rua Cumberland nº 550, sala 154, Ottawa, Ontario, K1N 6N5, 613-562-5387, ou pelo email ethics@uottwa.ca.

Há duas cópias deste formulário de consentimento, uma delas é sua.

Participante

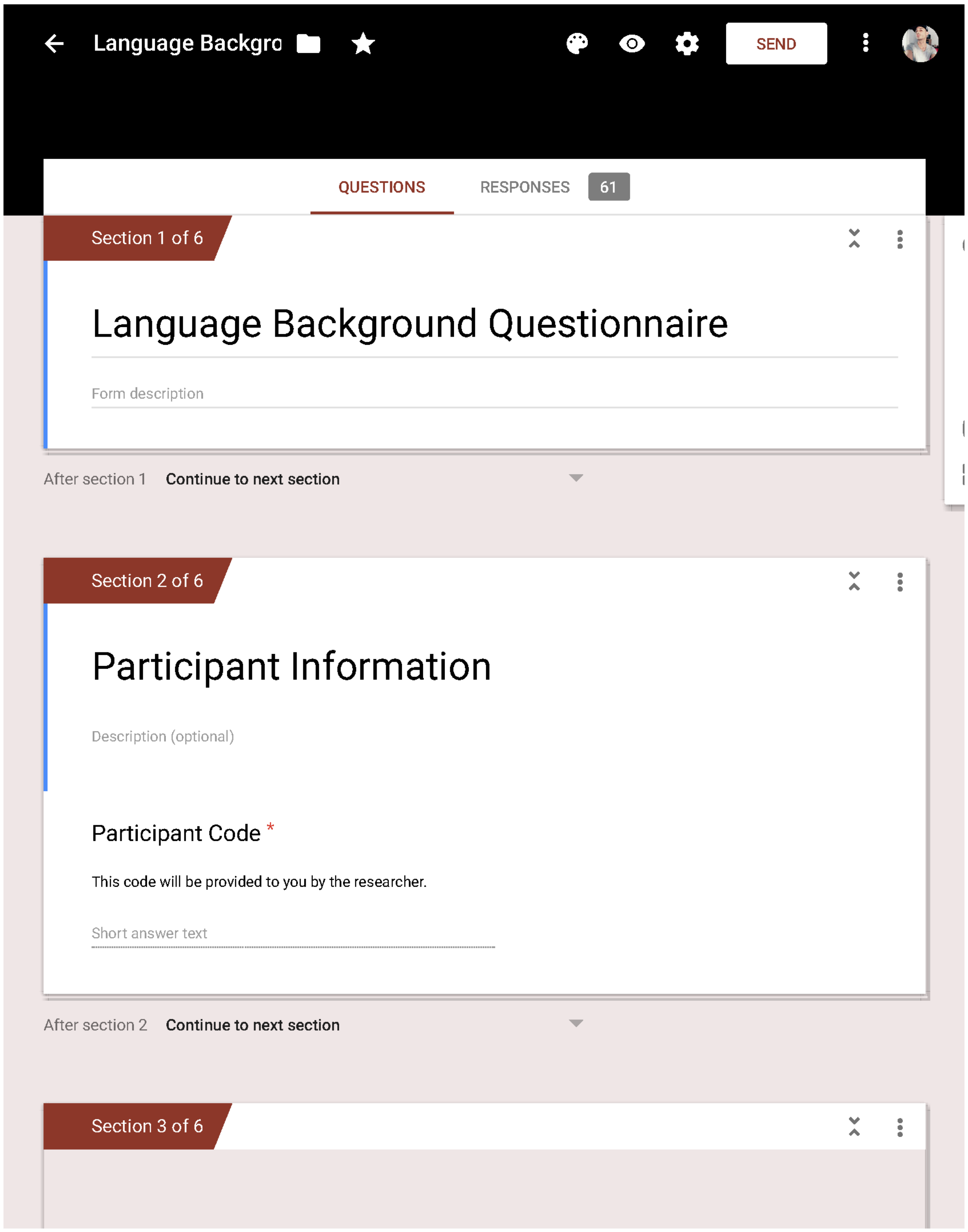
Nome Assinatura Data

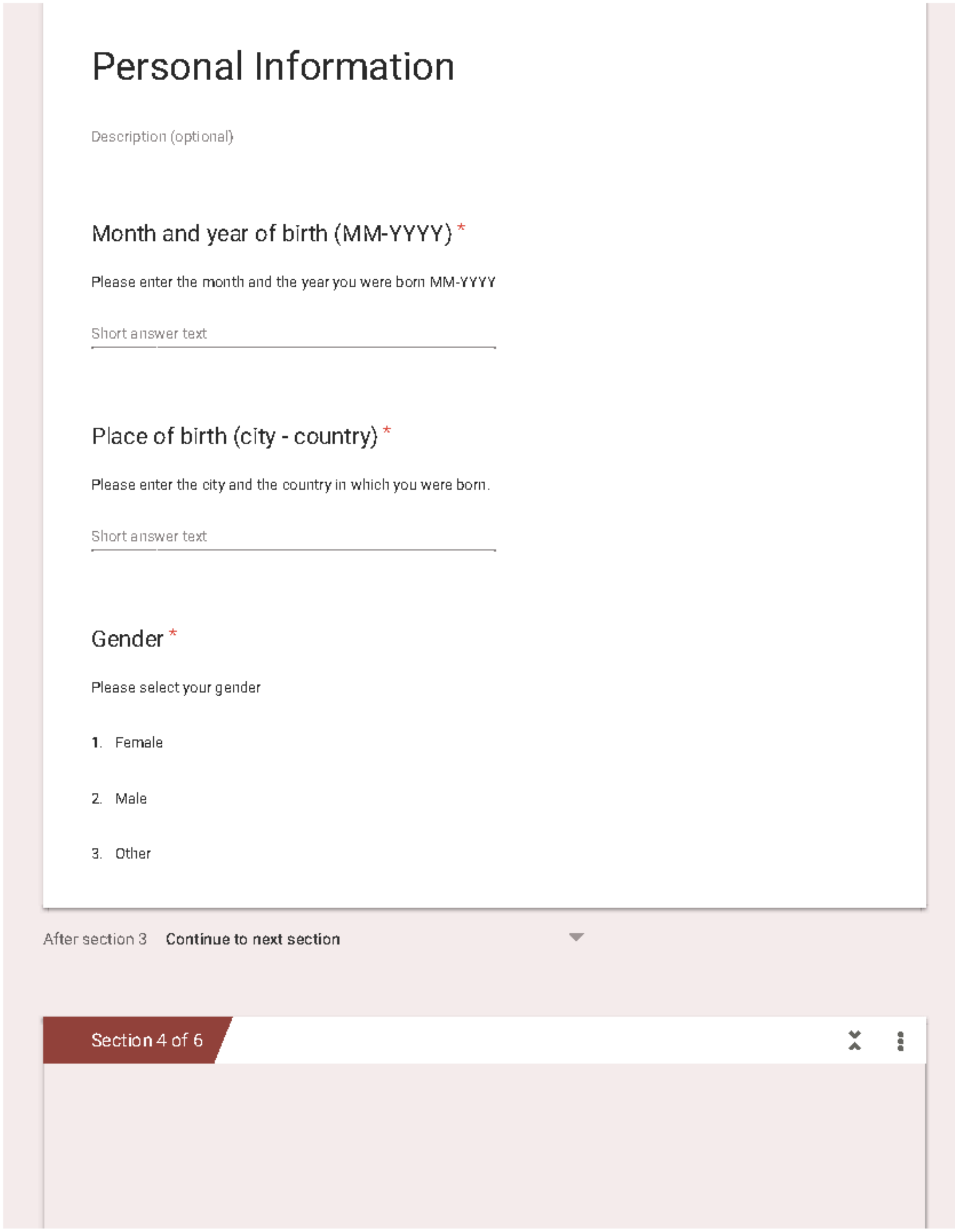
Investigador Principal

Nome Assinatura Data

613-562-5286
613-562-5141
70 Laurier E.
Ottawa ON K1N 6N5 Canada
www.uOttawa.ca

Appendix D: Language Background Questionnaire





Personal Information

Description (optional)

Month and year of birth (MM-YYYY) *

Please enter the month and the year you were born MM-YYYY

Short answer text

Place of birth (city - country) *

Please enter the city and the country in which you were born.

Short answer text

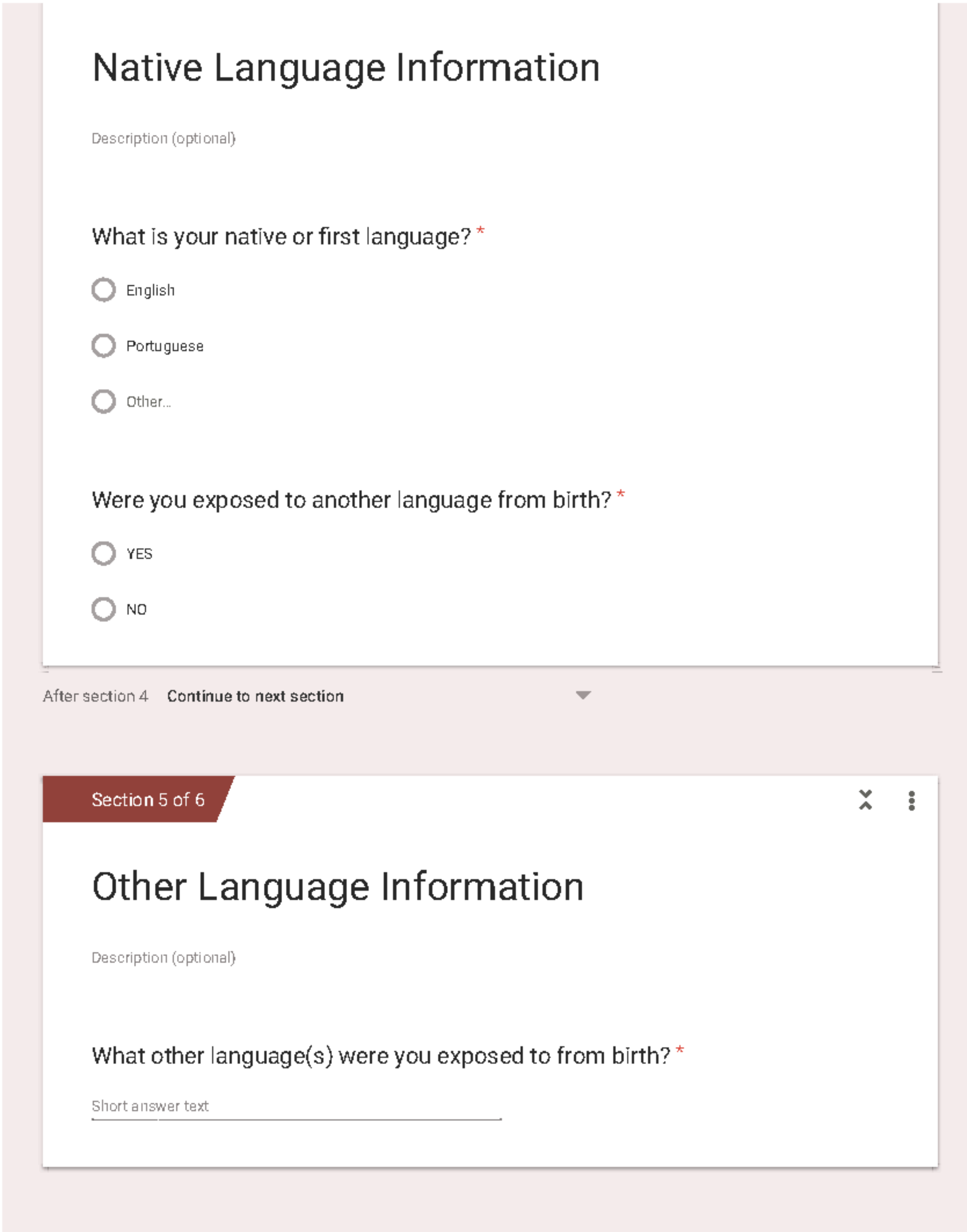
Gender *

Please select your gender

1. Female
2. Male
3. Other

After section 3 Continue to next section

Section 4 of 6



Native Language Information

Description (optional)

What is your native or first language? *

- English
- Portuguese
- Other...

Were you exposed to another language from birth? *

- YES
- NO

After section 4 Continue to next section

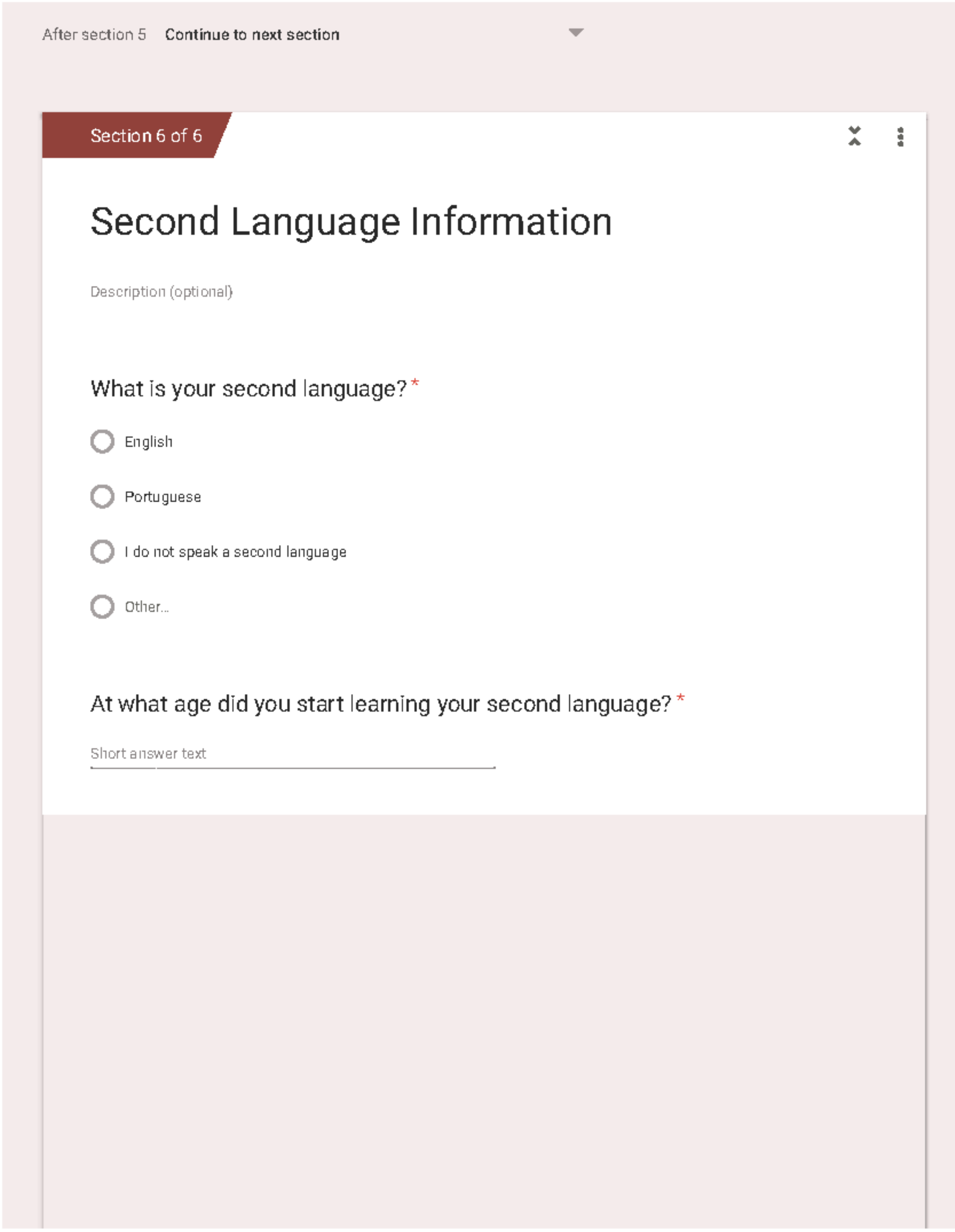
Section 5 of 6

Other Language Information

Description (optional)

What other language(s) were you exposed to from birth? *

Short answer text



After section 5 Continue to next section

Section 6 of 6

Second Language Information

Description (optional)

What is your second language? *

- English
- Portuguese
- I do not speak a second language
- Other...

At what age did you start learning your second language? *

Short answer text

Where did you learn your second language? *

- ○ elementary school
- ○ high school
- ○ college/university
- ○ language school
- ○ family (parents or relatives)
- ○ Other...

How often do you use your second language now? *

- ○ everyday
- ○ very often
- ○ often
- ○ sometimes
- ○ rarely
- ○ never

On the scale below, what is your overall level of proficiency in your second language? *

	1	2	3	4	5	6	7	8	9	10	
low proficiency	○	○	○	○	○	○	○	○	○	○	high proficiency

On the scale below, please rate your level of writing proficiency in your second language? *

	1	2	3	4	5	6	7	8	9	10	
low proficiency	○	○	○	○	○	○	○	○	○	○	high proficiency

On the scale below, please rate your level of reading proficiency in your second language? *

	1	2	3	4	5	6	7	8	9	10	
low proficiency	○	○	○	○	○	○	○	○	○	○	high proficiency

On the scale below, please rate your level of oral proficiency in your second language? *

	1	2	3	4	5	6	7	8	9	10	
low proficiency	○	○	○	○	○	○	○	○	○	○	high proficiency

Appendix E: Cloze Proficiency Test

Participant Code: ____________________

Cloze Test (Brown, 1980)

DIRECTIONS

1. Read the passage quickly to get the general meaning.
2. Write only one word in each blank next to the item number.
3. Contractions such as ***don't*** and ***John's*** bicycle are one word.
4. Check your answers.
5. Spelling will not count against you as long as the words in the blanks are legible.

EXAMPLE
The boy walked up the street. He stepped on a piece of ice. He fell (1) ***down*** , but he didn't hurt himself.

MAN AND HIS PROGRESS

Man is the only living creature that can make and use tools. He is the most teachable of living beings, earning the name of Homo sapiens. (1)_______________ever restless brain has used the 2)_______________and the wisdom of his ancestors (3)_____________ improve his way of life. Since (4)_____________ is able to walk and run (5)_______________ his feet, his hands have always (6)_______________free to carry and to use (7)_______________. Man's hands have served him well (8) ______________ his life on earth. His development, (9)______________can be divided into three major (10) _______________, is marked by several different ways (11) _______________life.

Up to 10,000 years ago, (12) _______________ human beings lived by hunting and (13) _______________. They also picked berries and fruits, (14) _______________ dug for various edible roots. Most (15) ______________, the men were the hunters, and (16) _______________ women acted as food gatherers. Since (17) _______________women were busy with the children, (18) _______________men handled the tools. In a

(19) ________________ hand, a dead branch became a (20) ________________ to knock down fruit or to (21) ________________ for tasty roots. Sometimes, an animal (22) ________________ served as a club, and a (23)________________ piece of stone, fitting comfortably into (24) ________________ hand, could be used to break (25)____________ or to throw at an animal. (26) ________________ stone was chipped against another until (27) ________________had a sharp edge. The primitive (28)________________ who first thought of putting a (29) ________________ stone at the end of a (30) ________________ made a brilliant discovery: he (31) ________________ joined two things to make a (32) ________________ useful tool, the spear. Flint, found (33) ________________ many rocks, became a common cutting (34) ________________ in the Paleolithic period of man's (35) ________________. Since no wood or bone tools (36) ________________ survived, we know of this man (37)_______________ his stone implements, with which he (38) ________________ kill animals, cut up the meat, (39) ________________ scrape the skins, as well as (40) _____________ pictures on the walls of the (41)_____________ where he lived during the winter.

(42) ________________ the warmer seasons, man wandered on (43) ________________ steppes of Europe without a fixed (44) _______________, always foraging for food. Perhaps the (45)________________ carried nuts and berries in shells (46) __________ skins or even in light, woven (47) ________________. Wherever they camped, the primitive people (48) ________________fires by striking flint for sparks (49) ________________ using dried seeds, moss, and rotten (50) ________________ for tinder. With fires that he kindled himself, man could keep wild animals away and could cook those that he killed, as well as provide warmth and light for himself.

Appendix F: Experimental Word Lists

C: cognate; F: false friend; D: distractor/pseudoword; P: practice; CTRL: control

Word List 1

prime	TARGET	Exp Status	Item Nbr
abate	PRESS	CTRL	2
abrigo	ADOPTIVE	CTRL	3
acarajé	ALARM	CTRL	4
açude	ALCOHOL	CTRL	6
adesivo	ALIGNMENT	CTRL	7
agulha	ANGLE	CTRL	8
aipo	SIGNATURE	CTRL	9
alagamento	ASTUTE	CTRL	11
alcatrão	BAGGAGE	CTRL	12
alecrim	BAND	CTRL	13
alho	RIM	CTRL	14
alvenaria	BASE	CTRL	16
ameaça	SMOKING	CTRL	18
amigo	BATTERY	CTRL	19
anão	BEIGE	CTRL	20
anca	BOMB	CTRL	21
angu	BRAVE	CTRL	23
apoio	CAPACITY	CTRL	26
arnica	CARGO	CTRL	27
avestruz	CAR	CTRL	28
baba	CHEQUE	CTRL	30
baderna	CHOCOLATE	CTRL	31
bando	COMA	CTRL	34
baque	COMATOSE	CTRL	35
barata	COMEDY	CTRL	36
batente	COMPUTER	CTRL	38
berço	CONFIDENT	CTRL	40
bilhete	COMFORT	CTRL	41
boi	CONVENIENT	CTRL	43
bolo	CREAM	CTRL	45
cama	PULL	CTRL	51
caminhão	DETERGENT	CTRL	53
capuz	DIFFERENT	CTRL	54
carne	DISTRIBUTION	CTRL	55
carrapato	ELEVATOR	CTRL	57
cartomante	ENTRANCE	CTRL	58
cebola	EXCELLENT	CTRL	60
cílio	FATALITY	CTRL	63
cisco	FEVER	CTRL	64
ciúme	FECES	CTRL	65

Word List 1

prime	TARGET	Exp Status	Item Nbr
cochilo	FLEXIBLE	CTRL	66
congelador	SYMPATHETIC	CTRL	68
corneta	FRUIT	CTRL	71
couve	GEL	CTRL	73
criança	GENIAL	CTRL	75
cúmulo	HOUR	CTRL	78
curral	INGENUITY	CTRL	79
dano	SORT	CTRL	81
dardo	TAMPON	CTRL	82
débil	INTERPRETATION	CTRL	83
delito	LICENSE	CTRL	86
demão	LANGUAGE	CTRL	87
demasia	LINE	CTRL	88
dente	MANNEQUIN	CTRL	89
destroço	ULTIMATELY	CTRL	92
égua	RANGER	CTRL	94
elo	MINERAL	CTRL	96
fantoche	MOUSE	CTRL	102
farda	HASTE	CTRL	103
fenda	NATURE	CTRL	105
frei	SUM	CTRL	110
freio	TAX	CTRL	111
furão	TENANT	CTRL	114
gambá	OCEAN	CTRL	116
golpe	OTHER	CTRL	119
granada	PARTICULAR	CTRL	124
gravação	PIANO	CTRL	126
ida	VENT	CTRL	132
isca	PROPER	CTRL	134
jabuti	PSYCHOLOGY	CTRL	135
jaca	PSYCHOSIS	CTRL	136
jararaca	PUMA	CTRL	138
joaninha	RAT	CTRL	142
labirinto	REAL	CTRL	145
laca	REFERENCE	CTRL	146
ladra	REMEDY	CTRL	148
laguna	REPAIR	CTRL	149
laje	REQUIREMENT	CTRL	150
libélula	ROSE	CTRL	156
limão	RUDE	CTRL	157

Word List 1

prime	TARGET	Exp Status	Item Nbr
limo	RUMOR	CTRL	158
linho	SAMBA	CTRL	159
liquidificador	SENTIMENT	CTRL	160
lixo	SIGNIFICANCE	CTRL	162
loção	SIMPLE	CTRL	164
lona	TEMPERATURE	CTRL	167
louco	TOMATO	CTRL	170
manchete	TRANSFORMATION	CTRL	172
manicure	TREPIDATION	CTRL	174
manobra	TRIUMPH	CTRL	175
marimbondo	ULTIMATUM	CTRL	177
meado	VEGETABLE	CTRL	179
meia	VEIN	CTRL	180
melancia	VEHICLE	CTRL	181
metade	RECIPIENT	CTRL	183
metido	VINEGAR	CTRL	184
multa	VENGEANCE	CTRL	186
musculação	VOWEL	CTRL	188
nada	VOTE	CTRL	189
nau	ZEBRA	CTRL	190
obra	ARM	CTRL	197
olaria	ASSIST	CTRL	198
ombro	ACTUAL	CTRL	199
onça	AUDIENCE	CTRL	200
paiol	BANK	CTRL	203
papo	BEEF	CTRL	205
paquiderme	BOND	CTRL	206
pipa	COLLAR	CTRL	217
prata	COMPROMISE	CTRL	221
preguiça	CONCOURSE	CTRL	223
prenda	CONCEAL	CTRL	224
propina	COSTUME	CTRL	226
punho	CUSP	CTRL	229
purê	DATA	CTRL	230
quadra	DESSERT	CTRL	231
raio	ENJOY	CTRL	237
ralo	EXPERT	CTRL	238
ramo	EXQUISITE	CTRL	239
rapaz	ORDINARY	CTRL	292
rena	EXCITING	CTRL	242

Word List 1

prime	TARGET	Exp Status	Item Nbr
renda	EXIT	CTRL	243
réu	FACILITY	CTRL	246
roda	GRIP	CTRL	249
ruga	HOME	CTRL	251
saguão	HOSPICE	CTRL	252
saia	IDIOM	CTRL	253
saldo	INSCRIPTION	CTRL	256
sarda	INGENIOUS	CTRL	254
sinal	INSTANCE	CTRL	257
sobra	JAR	CTRL	258
sobrado	JOURNAL	CTRL	259
solteiro	LACE	CTRL	260
soluço	LAMP	CTRL	261
sortudo	RECOURSE	CTRL	263
sujeira	LASER	CTRL	264
toldo	LENS	CTRL	267
trilho	LOCATION	CTRL	269
tristeza	LUNCH	CTRL	270
unha	MAGAZINE	CTRL	271
urubu	MASCARA	CTRL	273
velho	MALL	CTRL	275
velocímetro	MASTER	CTRL	276
verruga	NOTICE	CTRL	278
véu	OUTDOOR	CTRL	281
visom	PAVEMENT	CTRL	287
vovó	POST	CTRL	289
voz	PREJUDICE	CTRL	291
aba	HEVA	D	1
abestalhado	CHECKLEBURST	D	2
açúcar	HARETS	D	3
adorno	NULIPS	D	4
agrião	QUELTER	D	5
água	CAHED	D	6
alavanca	BETTULO	D	7
alcatra	VALLUME	D	227
alergia	OFFEN	D	8
alfazema	HOOPLES	D	9
aluguel	SUNTELO	D	10
amaciante	FRONG	D	11
amendoim	BELISM	D	12

Word List 1

prime	TARGET	Exp Status	Item Nbr
andaime	SEGRETS	D	13
arco-íris	OUTBRAW	D	14
asno	OLEALD	D	266
ata	KIFF	D	15
aterrisagem	PAMA	D	16
azeitona	CUNWREMY	D	17
babosa	WINOOZE	D	275
bafo	VEAST	D	18
bagulho	REOCCUS	D	19
bandeja	PUNOUN	D	20
banguela	MALATROSE	D	21
banheira	INWORM	D	22
baranga	MUNPY	D	23
barco	OVENCY	D	24
barro	ESUYP	D	284
barulho	RERARD	D	25
baunilha	RERATE	D	26
berimbau	BESLOOR	D	253
bina	TOOGIT	D	282
biscate	CHORENN	D	270
bocejo	VOQUEV	D	267
bofe	MICRONY	D	27
bofetada	TRONSHAT	D	28
borracheiro	MOGININ	D	29
brega	BETIPE	D	30
bruxa	HOLS	D	31
buço	VOOV	D	226
caçamba	CHEAHOOF	D	242
cachaça	NUCKE	D	32
cafetão	SWOPOM	D	241
cafofo	PLOCK	D	33
cafuné	HUNM	D	34
calçada	CHEIR	D	35
camada	MUEBA	D	36
câmbio	MERIP	D	37
caneca	UNELIND	D	235
caneta	LADLOP	D	38
cangote	MALATED	D	39
capim	BELYCH	D	40
capô	CHEESTNAIR	D	41

Word List 1

prime	TARGET	Exp Status	Item Nbr
caraca	JIOFRAX	D	246
carango	TINEREW	D	276
careca	MODGONE	D	237
carioca	WERRADITH	D	271
caroço	RODRUP	D	278
carrasco	KARPLE	D	42
carreta	TELCI	D	43
carroça	MORATH	D	279
casa	PIFFURE	D	44
cascavel	INTIENT	D	45
caxumba	DRIRATHIEL	D	245
cenoura	BEORARNI	D	46
cera	KREEPEE	D	47
chácara	UNDERDOUG	D	248
chacrete	TOXIMBLE	D	264
chafariz	POLLNEST	D	48
chão	SOTTA	D	49
chicória	ODRANOEL	D	50
chinelo	OPERTOT	D	51
chuchu	SHABI	D	52
cidadão	ENFLEP	D	53
cinto	REVENTIAL	D	54
coisa	TEBS	D	55
coleira	JOHACKLE	D	236
colete	ADNANREF	D	56
comício	JOVAPHILE	D	225
comida	PROFINROR	D	57
coqueluche	KULATOR	D	58
correio	HARPENITAGE	D	59
couro	FOMMER	D	60
coxinha	PLOOSNAR	D	230
cozinha	CHOOR	D	61
dendê	SUMNIKA	D	62
desemprego	FUTUKE	D	63
deus	WOTNOT	D	64
dragão	TESBAR	D	65
erva	ZESHO	D	66
esfregão	DARNET	D	67
esgoto	LEPSET	D	68
estrume	CHACAKA	D	247

Word List 1

prime	TARGET	Exp Status	Item Nbr
farol	RESAIX	D	281
fechadura	DINNY	D	69
feijoada	PROTHORACLED	D	70
ferro	JUNSET	D	71
fiel	SERAOS	D	72
foca	WENSH	D	73
fogo	BIEDE	D	74
fogueira	MBSON	D	75
framboesa	RELONE	D	76
frigideira	KECHIN	D	77
fronha	CHURRUSTUL	D	78
frouxo	WALNERT	D	79
furacão	TRINETTE	D	80
furo	TWM	D	294
gaita	NIETWE	D	277
garfo	PLAMPLE	D	81
gentalha	BROXNER	D	82
geringonça	MONTAQUEST	D	83
girico	GROJET	D	84
goela	MIRESA	D	251
goma	NEKMIT	D	283
graveto	OBBISN	D	85
gravura	NOURINE	D	86
gula	BUEND	D	293
henê	JOURCH	D	87
hipoteca	JOENTELINITY	D	88
hóspede	FEA	D	89
imóvel	CAKESKI	D	90
impressora	DOCTURE	D	91
inhame	HEMICAL	D	92
invalidez	NEXTING	D	93
jacaré	PROUSE	D	94
janela	PUBIEN	D	95
jeca	TROSP	D	96
jejum	NONVICT	D	97
joia	CHIER	D	98
jumento	ADEMPT	D	99
lagarta	WITHNEN	D	100
lagartixa	MINSE	D	101
lapela	PORTHOLOBATE	D	102

Word List 1

prime	TARGET	Exp Status	Item Nbr
laranja	FROMMIN	D	103
lareira	ANUNAT	D	104
leite	MEGTA	D	105
lerdo	IBBLY	D	106
lesma	QUONBI	D	107
liminar	KROWT	D	108
lula	RACHLER	D	109
lustre	CUTION	D	110
luz	RELAND	D	111
macaco	CLORE	D	112
maçarico	LAWMER	D	113
macete	NUSHTRA	D	114
maço	PLINGT	D	115
macumba	PEEVISITION	D	116
madeixa	RERATE	D	117
mala	PARIDIK	D	118
malhação	JAYL	D	119
mamadeira	EDADIVITAN	D	120
mangueira	TRIME	D	121
manicômio	YONNAL	D	122
manifestante	LOGITUDE	D	123
maremoto	YENBATUR	D	124
marimba	SAFOME	D	239
matraca	CHEILITH	D	252
meleca	AFFER	D	125
merenda	GLARETRAM	D	262
mês	DURETS	D	126
mesa	GLANGING	D	127
meta	LACTER	D	128
mexerica	BUMMOXA	D	240
milho	JUESMA	D	129
minhoca	WANTEMER	D	130
moçoila	SASAROO	D	243
mocréia	NETTE	D	131
moeda	SUBAVA	D	132
moela	TITUI	D	133
mofo	ZOIT	D	134
moita	RUCTULE	D	135
molambento	MUFTA	D	139
molho	NEVESY	D	136

Word List 1

prime	TARGET	Exp Status	Item Nbr
monge	RACA	D	137
moqueca	CHERSHOEE	D	255
morango	CLICEZO	D	138
mudo	ECKLE	D	261
muleta	TUPACASE	D	254
mulher	APINGO	D	140
mundo	NYNDRA	D	141
muquifo	CUTLUS	D	142
mutreta	CATACELLA	D	143
nabo	RUITA	D	144
namoro	BETTARAN	D	145
nanico	ANIRAM	D	146
nariz	FUDE	D	147
neném	COMED	D	148
ninho	ZIBLA	D	149
óculos	SIEDCASE	D	150
olheira	SANNY	D	151
olho	DRAKET	D	152
ouvinte	RUXTA	D	153
parceiro	FLIPKNIPS	D	154
pasta	EVENULE	D	155
pata	TUPIB	D	288
pauta	MULTIVE	D	156
paviu	CHILLAID	D	263
peixe	FUEVE	D	157
pena	WASMAC	D	158
perícia	VIOTTIS	D	244
periguete	AFFERARON	D	268
peruca	BIASDO	D	257
picanha	PELOOZOID	D	228
pichação	OHCAMARG	D	159
pneu	HAPPER	D	160
pochete	NAPERONE	D	286
pocilga	LALONTENT	D	161
pomar	NECCA	D	162
praia	HYDRING	D	163
presunto	PADA	D	164
primavera	INHOFT	D	165
queda	LAUDEITNESS	D	166
quibe	BUMOLA	D	232

Word List 1

prime	TARGET	Exp Status	Item Nbr
quinze	PHICOLY	D	167
rabisco	GROLE	D	168
rabo	SCHNOL	D	169
rachadura	SPIPSCOPS	D	170
ranço	KERBAL	D	171
raposa	POMODO	D	172
redação	GULLSTONT	D	173
reduto	KUCALIR	D	174
régua	SADOKIN	D	175
relógio	ACAVO	D	176
remo	LANKERTORT	D	177
revogação	KROSHIST	D	178
rissole	PRUVIA	D	231
rodovia	KATASTRY	D	179
roleta	BOCILILE	D	249
rolha	OVERSAL	D	180
rótulo	PORIC	D	181
sabor	SETENFI	D	182
sacana	VASAGLE	D	292
saci	OSHUN	D	183
sacolé	TESHEKUR	D	184
salpicão	DIGISOL	D	234
salsinha	HIABBYA	D	185
sapeca	OLIELLE	D	238
sarro	BISTUP	D	287
sebo	JEREN	D	273
sede	YEAVE	D	186
serelepe	LOOPLAB	D	229
seta	ROINAD	D	233
sinuca	SEILIU	D	256
sol	ENENTINESS	D	187
surdo	YOFFA	D	260
sutiã	COSOMOQ	D	188
tabefe	SPUSIOUS	D	291
tamanduá	MARESOM	D	189
tapete	FROWING	D	190
tarado	QIAMETH	D	285
tatú	ERAOW	D	259
teclado	MOVINY	D	191
terremoto	AVILI	D	192

Word List 1

prime	TARGET	Exp Status	Item Nbr
testa	OPONT	D	193
tigela	ARIGAIT	D	194
tormenta	IMPLOCE	D	195
torresmo	FLIONDESO	D	250
toucinho	RANTOSTNER	D	196
traição	KONKERS	D	197
tranco	ACAER	D	272
trema	ALQUE	D	198
tribufu	HENDASSA	D	274
trinco	REILTAS	D	289
troco	SATEIL	D	290
troço	YABATA	D	199
tromba	MURER	D	200
trono	QUINK	D	201
trote	YIMELLO	D	280
truta	FRUGTER	D	202
tubarão	CONTERAB	D	203
tutu	UMPRE	D	204
vagão	ONAMA	D	269
varal	ELOCIN	D	205
varejo	COPIC	D	206
vazamento	ZUZUUT	D	207
vazão	SOAXCASE	D	208
vela	PUSOON	D	209
vencimento	CINTRERY	D	210
verão	YULANAR	D	211
vereda	DORITY	D	212
verniz	TAMTOP	D	213
vesgo	GORMET	D	214
vista	TELVIT	D	215
vitrine	HIPPLED	D	216
vitrô	BETIFT	D	217
vitrola	NOVALY	D	258
viveiro	ZEANDROT	D	218
xaveco	NOCOBOT	D	265
xilindró	GAULIN	D	219
xinxim	PHREST	D	220
xodó	ATINTE	D	221
zabumba	SHINESH	D	222
zagueiro	BRUXET	D	223

Word List 1

prime	TARGET	Exp Status	Item Nbr
zarolho	NENURIA	D	224
acento	ACCENT	F	1
adepto	ADEPT	F	2
adjunto	ADJUNCT	F	3
agenda	AGENDA	F	4
aplicação	APPLICATION	F	5
assiste	ASSIST	F	7
audiência	AUDIENCE	F	9
banco	BANK	F	12
bife	BEEF	F	14
brando	BRAND	F	17
carta	CART	F	19
cartão	CARTON	F	20
chope	CHOP	F	23
chute	CHUTE	F	24
colar	COLLAR	F	26
conceito	CONCEIT	F	31
cova	COVE	F	36
cuspe	CUSP	F	38
data	DATA	F	39
deserto	DESSERT	F	40
diversão	DIVERSION	F	42
draga	DRAG	F	43
editor	EDITOR	F	44
esperto	EXPERT	F	46
estrangeiro	ESTRANGED	F	48
estranho	STRANGER	F	49
êxito	EXIT	F	51
fábrica	FABRIC	F	53
facilidade	FACILITY	F	54
graduação	GRADUATION	F	55
gratuito	GRATUITY	F	56
inscrição	INSCRIPTION	F	65
lanche	LUNCH	F	77
legenda	LEGEND	F	72
locação	LOCATION	F	76
maior	MAYOR	F	79
máscara	MASCARA	F	80
mau	MALL	F	82
notícia	NOTICE	F	85

Word List 1

prime	TARGET	Exp Status	Item Nbr
novela	NOVEL	F	86
ofício	OFFICE	F	87
ordinário	ORDINARY	F	127
outdoor	OUTDOOR	F	88
pão	PAN	F	89
parente	PARENT	F	90
pastel	PASTEL	F	91
pavimento	PAVEMENT	F	93
pote	POT	F	96
prejuízo	PREJUDICE	F	97
preservativo	PRESERVATIVE	F	98
pressa	PRESS	F	99
pretende	PRETEND	F	100
recurso	RECOURSE	F	106
resumo	RÉSUMÉ	F	107
simpático	SYMPATHETIC	F	112
smoking	SMOKING	F	113
sorte	SORT	F	115
suporte	SUPPORT	F	117
tampão	TAMPON	F	118
taxa	TAX	F	119
tenente	TENANT	F	120
turno	TURN	F	123
vento	VENT	F	125
vila	VILE	F	126
abajur	BANDOVA	P	1
acre	DRIM	P	2
apito	WHISTLE	P	3
aranha	SHENDON	P	4
aumento	RAISE	P	5
balde	WOMT	P	6
batata	DONCE	P	7
bebida	BEVERAGE	P	8
broca	YOENK	P	9
caixote	KAHOE	P	10
campo	DIELB	P	11
carretel	REEL	P	12
cogumelo	MUSHROOM	P	13
crença	BELIEF	P	14
descarado	BLATANT	P	15

Word List 1

prime	TARGET	Exp Status	Item Nbr
desejo	URGE	P	16
enguia	EEL	P	17
ferrugem	RUST	P	18
gema	YOLK	P	19
irritado	LASATY	P	20
isopor	STYROFOAM	P	21
linguiça	SAUSAGE	P	22
luva	GLOVE	P	23
magricela	LANKY	P	24
manivela	CRANK	P	25
marinheiro	SITUYA	P	26
mobília	FURNITURE	P	27
óbito	PRINKT	P	28
oficina	HENTIOM	P	29
recife	REEF	P	30
sótão	ATTIC	P	31
tesoura	SCISSORS	P	32
tesoura	SCISSORS	P	32
touro	BULL	P	33
valsa	WALTZ	P	34
vaqueiro	COWBOY	P	35
abrupto	ABRUPT	C	1
adotivo	ADOPTIVE	C	2
álcool	ALCOHOL	C	4
ângulo	ANGLE	C	6
base	BASE	C	13
bateria	BATTERY	C	14
botão	BUTTON	C	17
capacidade	CAPACITY	C	21
caso	CASE	C	24
cheque	CHEQUE	C	25
coluna	COLUMN	C	28
comatoso	COMATOSE	C	30
comédia	COMEDY	C	31
complicado	COMPLICATED	C	32
computador	COMPUTER	C	33
condicional	CONDITIONAL	C	34
confidente	CONFIDENT	C	35
conforto	COMFORT	C	36
cortina	CURTAIN	C	39

Word List 1

prime	TARGET	Exp Status	Item Nbr
criticismo	CRITICISM	C	41
cubo	CUBE	C	42
dama	DAME	C	43
detergente	DETERGENT	C	47
diferente	DIFFERENT	C	48
distribuição	DISTRIBUTION	C	49
elástico	ELASTIC	C	50
elevador	ELEVATOR	C	51
entrada	ENTRANCE	C	52
estúpido	STUPID	C	53
excelente	EXCELLENT	C	54
fantasia	FANTASY	C	56
febre	FEVER	C	58
fezes	FECES	C	59
flexível	FLEXIBLE	C	60
frustrado	FRUSTRATED	C	63
fruta	FRUIT	C	64
gel	GEL	C	66
genial	GENIAL	C	68
girafa	GIRAFFE	C	69
manequim	MANNEQUIN	C	79
máquina	MACHINE	C	80
mineral	MINERAL	C	84
mosaico	MOSAIC	C	86
mosquito	MOSQUITO	C	87
mostarda	MUSTARD	C	88
museu	MUSEUM	C	91
natureza	NATURE	C	92
ocasião	OCCASION	C	96
outro	OTHER	C	100
paciência	PATIENCE	C	101
pagamento	PAYMENT	C	102
papel	PAPER	C	103
parque	PARK	C	104
piano	PIANO	C	107
planta	PLANT	C	108
prisão	PRISON	C	111
proficiência	PROFICIENCY	C	112
projetor	PROJECTOR	C	113
próprio	PROPER	C	114

Word List 1

prime	TARGET	Exp Status	Item Nbr
rato	RAT	C	120
referência	REFERENCE	C	123
reparo	REPAIR	C	126
resposta	RESPONSE	C	128
restaurante	RESTAURANT	C	129
ritmo	RHYTHM	C	131
rosa	ROSE	C	133
shampoo	SHAMPOO	C	138
sofá	SOFA	C	142
sucesso	SUCCESS	C	143
testamento	TESTAMENT	C	145
texto	TEXT	C	146
transformação	TRANSFORMATION	C	149
trepidação	TREPIDATION	C	151
ultimato	ULTIMATUM	C	154
vagabundo	VAGABOND	C	155
vegetal	VEGETABLE	C	156
veia	VEIN	C	157
verme	VERMIN	C	159
vinagre	VINEGAR	C	160
vingança	VENGEANCE	C	161
vocação	VOCATION	C	162
zebra	ZEBRA	C	165
zero	ZERO	C	166

Word List 2

prime	TARGET	Exp Status	Item Nbr
abalo	ABRUPT	CTRL	1
aceno	SERVICE	CTRL	5
ala	ASSISTANCE	CTRL	10
almoxarifado	BAR	CTRL	15
alvo	SAP	CTRL	17
anel	BUTTON	CTRL	22
animal	COFFEE	CTRL	24
anzol	CAMEL	CTRL	25
azulejo	CASE	CTRL	29
baldeação	SURGERY	CTRL	32
baleia	COLUMN	CTRL	33
batedeira	COMPLICATED	CTRL	37
beco	CONDITIONAL	CTRL	39
birra	CONSTIPATION	CTRL	42
boina	CURTAIN	CTRL	44
breu	CRITICISM	CTRL	46
brinco	CUBE	CTRL	47
bujão	DAME	CTRL	48
cabelo	DEMAND	CTRL	49
calça	DISGUST	CTRL	50
cambalhota	DISGRACE	CTRL	52
cara	ELASTIC	CTRL	56
caspa	STUPID	CTRL	59
chá	FACE	CTRL	61
ciente	FANTASY	CTRL	62
coice	FORMAT	CTRL	67
cor	FREEZER	CTRL	69
corja	FRUSTRATED	CTRL	70
corno	FUNDAMENTAL	CTRL	72
crachá	GELATIN	CTRL	74
crisma	GIRAFFE	CTRL	76
cueca	GRAPHITE	CTRL	77
dado	SOAP	CTRL	80
declive	IRATE	CTRL	84
dedo	LIBERTY	CTRL	85
derrota	MACHINE	CTRL	90
despacho	MEDAL	CTRL	91
dor	MEMORY	CTRL	93
eixo	METAL	CTRL	95

Word List 2

prime	**TARGET**	**Exp Status**	**Item Nbr**
embalagem	MONSTER	CTRL	97
escada	MOSAIC	CTRL	98
escama	MOSQUITO	CTRL	99
espantalho	MUSTARD	CTRL	100
falcatrua	MOTORCYCLE	CTRL	101
fatia	MUSEUM	CTRL	104
feriado	NEGOTIATE	CTRL	106
férias	NORMAL	CTRL	107
folha	TURN	CTRL	108
forno	NUMBER	CTRL	109
fuligem	OCCASION	CTRL	112
fundo	TOQUE	CTRL	113
gaiola	TERRACE	CTRL	115
garrafa	OIL	CTRL	117
gaveta	OPTION	CTRL	118
gosma	PATIENCE	CTRL	120
gota	PAYMENT	CTRL	121
goteira	PAPER	CTRL	122
grampo	PARK	CTRL	123
granel	PERSONAL	CTRL	125
greve	PLANT	CTRL	127
grife	PRACTICE	CTRL	128
grilo	PRECARIOUS	CTRL	129
grinalda	PRISON	CTRL	130
grito	PROFICIENCY	CTRL	131
infância	PROJECTOR	CTRL	133
janta	VILE	CTRL	137
jarro	QUANTITY	CTRL	139
jaula	RECEIPT	CTRL	140
jeito	RADIO	CTRL	141
joelho	REASON	CTRL	143
jovem	PRETEND	CTRL	144
lado	RELATIONSHIP	CTRL	147
lápis	RESPONSE	CTRL	151
laudo	RESTAURANT	CTRL	152
lavabo	RHYME	CTRL	153
lavadeira	RHYTHM	CTRL	154
lavradio	ROBOT	CTRL	155
lírio	SHAMPOO	CTRL	161

Word List 2

prime	TARGET	Exp Status	Item Nbr
lobo	SIMILAR	CTRL	163
lodo	SOFA	CTRL	165
lombriga	SUCCESS	CTRL	166
lontra	TESTAMENT	CTRL	168
lote	TEXT	CTRL	169
malandro	TRANSLATION	CTRL	171
mané	TRANSIT	CTRL	173
margarida	TROPHY	CTRL	176
martelo	VAGABOND	CTRL	178
merluza	VERMIN	CTRL	182
mosca	PUSH	CTRL	185
muro	VOCATION	CTRL	187
navalha	ZERO	CTRL	191
navio	ACCENT	CTRL	192
neblina	ADEPT	CTRL	193
negócio	ADJUNCT	CTRL	194
neve	AGENDA	CTRL	195
nexo	APPLICATION	CTRL	196
onda	HAZARD	CTRL	201
padeiro	BALCONY	CTRL	202
palito	BATON	CTRL	204
paquita	BRASS	CTRL	207
parapeito	BRAND	CTRL	208
pato	COMPASS	CTRL	209
pedra	CART	CTRL	210
pegada	CARTON	CTRL	211
peixaria	CHAPEL	CTRL	212
perdiz	CHAT	CTRL	213
pesquisa	CHOP	CTRL	214
piada	CHUTE	CTRL	215
pilantra	CLIQUE	CTRL	216
pirulito	COLLEGE	CTRL	218
pomada	COMPETITION	CTRL	219
ponte	COMPREHENSIBLE	CTRL	220
pregão	CONCEIT	CTRL	222
privada	CUP	CTRL	225
prostíbulo	COVE	CTRL	227
pulga	CURT	CTRL	228
quente	DISCUSSION	CTRL	232

Word List 2

prime	TARGET	Exp Status	Item Nbr
quiabo	DIVERSION	CTRL	233
quitute	DRAG	CTRL	234
rabanete	SUPPORT	CTRL	235
racha	EDITOR	CTRL	236
raquete	ESTRANGED	CTRL	240
rastro	STRANGER	CTRL	241
requinte	FAN	CTRL	244
reto	FABRIC	CTRL	245
risco	GRADUATION	CTRL	247
riso	GRATUITY	CTRL	248
ronco	DUEL	CTRL	293
rosca	GROCERY	CTRL	250
sarna	INJURY	CTRL	255
soneca	RÉSUMÉ	CTRL	262
talher	LEGEND	CTRL	265
tela	LECTURE	CTRL	266
trenó	LIMP	CTRL	268
urtiga	MAYOR	CTRL	272
vapor	MASS	CTRL	274
verde	MOISTURE	CTRL	277
vespa	NOVEL	CTRL	279
vestido	OFFICE	CTRL	280
viagem	PAN	CTRL	282
víbora	PARENT	CTRL	283
videira	RETIRED	CTRL	284
vidro	PASTEL	CTRL	285
virilha	PATRON	CTRL	286
voraz	POLICY	CTRL	288
vovô	POT	CTRL	290
zonzo	PRESERVATIVE	CTRL	292
aba	HEVA	D	1
abestalhado	CHECKLEBURST	D	2
açúcar	HARETS	D	3
adorno	NULIPS	D	4
agrião	QUELTER	D	5
água	CAHED	D	6
alavanca	BETTULO	D	7
alcatra	VALLUME	D	227
alergia	OFFEN	D	8

Word List 2

prime	TARGET	Exp Status	Item Nbr
alfazema	HOOPLES	D	9
aluguel	SUNTELO	D	10
amaciante	FRONG	D	11
amendoim	BELISM	D	12
andaime	SEGRETS	D	13
arco-íris	OUTBRAW	D	14
asno	OLEALD	D	266
ata	KIFF	D	15
aterrisagem	PAMA	D	16
azeitona	CUNWREMY	D	17
babosa	WINOOZE	D	275
bafo	VEAST	D	18
bagulho	REOCCUS	D	19
bandeja	PUNOUN	D	20
banguela	MALATROSE	D	21
banheira	INWORM	D	22
baranga	MUNPY	D	23
barco	OVENCY	D	24
barro	ESUYP	D	284
barulho	RERARD	D	25
baunilha	RERATE	D	26
berimbau	BESLOOR	D	253
bina	TOOGIT	D	282
biscate	CHORENN	D	270
bocejo	VOQUEV	D	267
bofe	MICRONY	D	27
bofetada	TRONSHAT	D	28
borracheiro	MOGININ	D	29
brega	BETIPE	D	30
bruxa	HOLS	D	31
buço	VOOV	D	226
caçamba	CHEAHOOF	D	242
cachaça	NUCKE	D	32
cafetão	SWOPOM	D	241
cafofo	PLOCK	D	33
cafuné	HUNM	D	34
calçada	CHEIR	D	35
camada	MUEBA	D	36
câmbio	MERIP	D	37

Word List 2

prime	TARGET	Exp Status	Item Nbr
caneca	UNELIND	D	235
caneta	LADLOP	D	38
cangote	MALATED	D	39
capim	BELYCH	D	40
capô	CHEESTNAIR	D	41
caraca	JIOFRAX	D	246
carango	TINEREW	D	276
careca	MODGONE	D	237
carioca	WERRADITH	D	271
caroço	RODRUP	D	278
carrasco	KARPLE	D	42
carreta	TELCI	D	43
carroça	MORATH	D	279
casa	PIFFURE	D	44
cascavel	INTIENT	D	45
caxumba	DRIRATHIEL	D	245
cenoura	BEORARNI	D	46
cera	KREEPEE	D	47
chácara	UNDERDOUG	D	248
chacrete	TOXIMBLE	D	264
chafariz	POLLNEST	D	48
chão	SOTTA	D	49
chicória	ODRANOEL	D	50
chinelo	OPERTOT	D	51
chuchu	SHABI	D	52
cidadão	ENFLEP	D	53
cinto	REVENTIAL	D	54
coisa	TEBS	D	55
coleira	JOHACKLE	D	236
colete	ADNANREF	D	56
comício	JOVAPHILE	D	225
comida	PROFINROR	D	57
coqueluche	KULATOR	D	58
correio	HARPENITAGE	D	59
couro	FOMMER	D	60
coxinha	PLOOSNAR	D	230
cozinha	CHOOR	D	61
dendê	SUMNIKA	D	62
desemprego	FUTUKE	D	63

Word List 2

prime	TARGET	Exp Status	Item Nbr
deus	WOTNOT	D	64
dragão	TESBAR	D	65
erva	ZESHO	D	66
esfregão	DARNET	D	67
esgoto	LEPSET	D	68
estrume	CHACAKA	D	247
farol	RESAIX	D	281
fechadura	DINNY	D	69
feijoada	PROTHORACLED	D	70
ferro	JUNSET	D	71
fiel	SERAOS	D	72
foca	WENSH	D	73
fogo	BIEDE	D	74
fogueira	MBSON	D	75
framboesa	RELONE	D	76
frigideira	KECHIN	D	77
fronha	CHURRUSTUL	D	78
frouxo	WALNERT	D	79
furacão	TRINETTE	D	80
furo	TWM	D	294
gaita	NIETWE	D	277
garfo	PLAMPLE	D	81
gentalha	BROXNER	D	82
geringonça	MONTAQUEST	D	83
girico	GROJET	D	84
goela	MIRESA	D	251
goma	NEKMIT	D	283
graveto	OBBISN	D	85
gravura	NOURINE	D	86
gula	BUEND	D	293
henê	JOURCH	D	87
hipoteca	JOENTELINITY	D	88
hóspede	FEA	D	89
imóvel	CAKESKI	D	90
impressora	DOCTURE	D	91
inhame	HEMICAL	D	92
invalidez	NEXTING	D	93
jacaré	PROUSE	D	94
janela	PUBIEN	D	95

Word List 2

prime	TARGET	Exp Status	Item Nbr
jeca	TROSP	D	96
jejum	NONVICT	D	97
joia	CHIER	D	98
jumento	ADEMPT	D	99
lagarta	WITHNEN	D	100
lagartixa	MINSE	D	101
lapela	PORTHOLOBATE	D	102
laranja	FROMMIN	D	103
lareira	ANUNAT	D	104
leite	MEGTA	D	105
lerdo	IBBLY	D	106
lesma	QUONBI	D	107
liminar	KROWT	D	108
lula	RACHLER	D	109
lustre	CUTION	D	110
luz	RELAND	D	111
macaco	CLORE	D	112
maçarico	LAWMER	D	113
macete	NUSHTRA	D	114
maço	PLINGT	D	115
macumba	PEEVISITION	D	116
madeixa	RERATE	D	117
mala	PARIDIK	D	118
malhação	JAYL	D	119
mamadeira	EDADIVITAN	D	120
mangueira	TRIME	D	121
manicômio	YONNAL	D	122
manifestante	LOGITUDE	D	123
maremoto	YENBATUR	D	124
marimba	SAFOME	D	239
matraca	CHEILITH	D	252
meleca	AFFER	D	125
merenda	GLARETRAM	D	262
mês	DURETS	D	126
mesa	GLANGING	D	127
meta	LACTER	D	128
mexerica	BUMMOXA	D	240
milho	JUESMA	D	129
minhoca	WANTEMER	D	130

Word List 2

prime	TARGET	Exp Status	Item Nbr
moçoila	SASAROO	D	243
mocréia	NETTE	D	131
moeda	SUBAVA	D	132
moela	TITUI	D	133
mofo	ZOIT	D	134
moita	RUCTULE	D	135
molambento	MUFTA	D	139
molho	NEVESY	D	136
monge	RACA	D	137
moqueca	CHERSHOEE	D	255
morango	CLICEZO	D	138
mudo	ECKLE	D	261
muleta	TUPACASE	D	254
mulher	APINGO	D	140
mundo	NYNDRA	D	141
muquifo	CUTLUS	D	142
mutreta	CATACELLA	D	143
nabo	RUITA	D	144
namoro	BETTARAN	D	145
nanico	ANIRAM	D	146
nariz	FUDE	D	147
neném	COMED	D	148
ninho	ZIBLA	D	149
óculos	SIEDCASE	D	150
olheira	SANNY	D	151
olho	DRAKET	D	152
ouvinte	RUXTA	D	153
parceiro	FLIPKNIPS	D	154
pasta	EVENULE	D	155
pata	TUPIB	D	288
pauta	MULTIVE	D	156
paviu	CHILLAID	D	263
peixe	FUEVE	D	157
pena	WASMAC	D	158
perícia	VIOTTIS	D	244
periguete	AFFERARON	D	268
peruca	BIASDO	D	257
picanha	PELOOZOID	D	228
pichação	OHCAMARG	D	159

Word List 2

prime	TARGET	Exp Status	Item Nbr
pneu	HAPPER	D	160
pochete	NAPERONE	D	286
pocilga	LALONTENT	D	161
pomar	NECCA	D	162
praia	HYDRING	D	163
presunto	PADA	D	164
primavera	INHOFT	D	165
queda	LAUDEITNESS	D	166
quibe	BUMOLA	D	232
quinze	PHICOLY	D	167
rabisco	GROLE	D	168
rabo	SCHNOL	D	169
rachadura	SPIPSCOPS	D	170
ranço	KERBAL	D	171
raposa	POMODO	D	172
redação	GULLSTONT	D	173
reduto	KUCALIR	D	174
régua	SADOKIN	D	175
relógio	ACAVO	D	176
remo	LANKERTORT	D	177
revogação	KROSHIST	D	178
rissole	PRUVIA	D	231
rodovia	KATASTRY	D	179
roleta	BOCILILE	D	249
rolha	OVERSAL	D	180
rótulo	PORIC	D	181
sabor	SETENFI	D	182
sacana	VASAGLE	D	292
saci	OSHUN	D	183
sacolé	TESHEKUR	D	184
salpicão	DIGISOL	D	234
salsinha	HIABBYA	D	185
sapeca	OLIELLE	D	238
sarro	BISTUP	D	287
sebo	JEREN	D	273
sede	YEAVE	D	186
serelepe	LOOPLAB	D	229
seta	ROINAD	D	233
sinuca	SEILIU	D	256

Word List 2

prime	TARGET	Exp Status	Item Nbr
sol	ENENTINESS	D	187
surdo	YOFFA	D	260
sutiã	COSOMOQ	D	188
tabefe	SPUSIOUS	D	291
tamanduá	MARESOM	D	189
tapete	FROWING	D	190
tarado	QIAMETH	D	285
tatú	ERAOW	D	259
teclado	MOVINY	D	191
terremoto	AVILI	D	192
testa	OPONT	D	193
tigela	ARIGAIT	D	194
tormenta	IMPLOCE	D	195
torresmo	FLIONDESO	D	250
toucinho	RANTOSTNER	D	196
traição	KONKERS	D	197
tranco	ACAER	D	272
trema	ALQUE	D	198
tribufu	HENDASSA	D	274
trinco	REILTAS	D	289
troco	SATEIL	D	290
troço	YABATA	D	199
tromba	MURER	D	200
trono	QUINK	D	201
trote	YIMELLO	D	280
truta	FRUGTER	D	202
tubarão	CONTERAB	D	203
tutu	UMPRE	D	204
vagão	ONAMA	D	269
varal	ELOCIN	D	205
varejo	COPIC	D	206
vazamento	ZUZUUT	D	207
vazão	SOAXCASE	D	208
vela	PUSOON	D	209
vencimento	CINTRERY	D	210
verão	YULANAR	D	211
vereda	DORITY	D	212
verniz	TAMTOP	D	213
vesgo	GORMET	D	214

Word List 2

prime	TARGET	Exp Status	Item Nbr
vista	TELVIT	D	215
vitrine	HIPPLED	D	216
vitrô	BETIFT	D	217
vitrola	NOVALY	D	258
viveiro	ZEANDROT	D	218
xaveco	NOCOBOT	D	265
xilindró	GAULIN	D	219
xinxim	PHREST	D	220
xodó	ATINTE	D	221
zabumba	SHINESH	D	222
zagueiro	BRUXET	D	223
zarolho	NENURIA	D	224
arma	ARM	F	6
atual	ACTUAL	F	8
azar	HAZARD	F	10
balcão	BALCONY	F	11
batom	BATON	F	13
bonde	BOND	F	15
braço	BRASS	F	16
bússola	COMPASS	F	18
chapéu	CHAPEL	F	21
chato	CHAT	F	22
clique	CLIQUE	F	25
colégio	COLLEGE	F	27
competição	COMPETITION	F	28
compreensível	COMPREHENSIBLE	F	29
compromisso	COMPROMISE	F	30
concurso	CONCOURSE	F	32
conselho	CONCEAL	F	33
copo	CUP	F	34
costume	COSTUME	F	35
curto	CURT	F	37
discussão	DISCUSSION	F	41
duelo	DUEL	F	128
enjoo	ENJOY	F	45
esquisito	EXQUISITE	F	47
excitante	EXCITING	F	50
fã	FAN	F	52
gripe	GRIP	F	57

Word List 2

prime	TARGET	Exp Status	Item Nbr
grosseria	GROCERY	F	58
haste	HASTE	F	59
homem	HOME	F	60
hospício	HOSPICE	F	61
idioma	IDIOM	F	62
ingênuo	INGENIOUS	F	63
injúria	INJURY	F	64
instância	INSTANCE	F	66
jarra	JAR	F	67
jornal	JOURNAL	F	68
laço	LACE	F	69
lâmpada	LAMP	F	70
lazer	LASER	F	71
leitura	LECTURE	F	73
lenço	LENS	F	74
limpo	LIMP	F	75
magazine	MAGAZINE	F	78
massa	MASS	F	81
mestre	MASTER	F	83
mistura	MOISTURE	F	84
patrão	PATRON	F	92
polícia	POLICY	F	94
poste	POST	F	95
pulo	PULL	F	101
puxe	PUSH	F	102
ranger	RANGER	F	103
receita	RECEIPT	F	104
recipiente	RECIPIENT	F	105
retirado	RETIRED	F	108
rim	RIM	F	109
sapo	SAP	F	110
serviço	SERVICE	F	111
sopa	SOAP	F	114
sumo	SUM	F	116
terraço	TERRACE	F	121
toque	TOQUE	F	122
ultimamente	ULTIMATELY	F	124
abajur	BANDOVA	P	1
acre	DRIM	P	2

Word List 2

prime	TARGET	Exp Status	Item Nbr
apito	WHISTLE	P	3
aranha	SHENDON	P	4
aumento	RAISE	P	5
balde	WOMT	P	6
batata	DONCE	P	7
bebida	BEVERAGE	P	8
broca	YOENK	P	9
caixote	KAHOE	P	10
campo	DIELB	P	11
carretel	REEL	P	12
cogumelo	MUSHROOM	P	13
crença	BELIEF	P	14
descarado	BLATANT	P	15
desejo	URGE	P	16
enguia	EEL	P	17
ferrugem	RUST	P	18
gema	YOLK	P	19
irritado	LASATY	P	20
isopor	STYROFOAM	P	21
linguiça	SAUSAGE	P	22
luva	GLOVE	P	23
magricela	LANKY	P	24
manivela	CRANK	P	25
marinheiro	SITUYA	P	26
mobília	FURNITURE	P	27
óbito	PRINKT	P	28
oficina	HENTIOM	P	29
recife	REEF	P	30
sótão	ATTIC	P	31
tesoura	SCISSORS	P	32
tesoura	SCISSORS	P	32
touro	BULL	P	33
valsa	WALTZ	P	34
vaqueiro	COWBOY	P	35
alarme	ALARM	C	3
alinhamento	ALIGNMENT	C	5
assinatura	SIGNATURE	C	7
assistência	ASSISTANCE	C	8
astuto	ASTUTE	C	9

Word List 2

prime	TARGET	Exp Status	Item Nbr
bagagem	BAGGAGE	C	10
banda	BAND	C	11
bar	BAR	C	12
bege	BEIGE	C	15
bomba	BOMB	C	16
bravo	BRAVE	C	18
café	COFFEE	C	19
camelo	CAMEL	C	20
cargo	CARGO	C	22
carro	CAR	C	23
chocolate	CHOCOLATE	C	26
cirurgia	SURGERY	C	27
coma	COMA	C	29
constipação	CONSTIPATION	C	37
conveniente	CONVENIENT	C	38
creme	CREAM	C	40
demanda	DEMAND	C	44
desgosto	DISGUST	C	45
desgraça	DISGRACE	C	46
face	FACE	C	55
fatalidade	FATALITY	C	57
formato	FORMAT	C	61
freezer	FREEZER	C	62
fundamental	FUNDAMENTAL	C	65
gelatina	GELATIN	C	67
grafite	GRAPHITE	C	70
hora	HOUR	C	71
ingenuidade	INGENUITY	C	72
interpretação	INTERPRETATION	C	73
irado	IRATE	C	74
liberdade	LIBERTY	C	75
licença	LICENSE	C	76
língua	LANGUAGE	C	77
linha	LINE	C	78
medalha	MEDAL	C	81
memória	MEMORY	C	82
metal	METAL	C	83
monstro	MONSTER	C	85
motocicleta	MOTORCYCLE	C	89

Word List 2

prime	TARGET	Exp Status	Item Nbr
mouse	MOUSE	C	90
negociar	NEGOTIATE	C	93
normal	NORMAL	C	94
número	NUMBER	C	95
oceano	OCEAN	C	97
óleo	OIL	C	98
opção	OPTION	C	99
particular	PARTICULAR	C	105
pessoal	PERSONAL	C	106
prática	PRACTICE	C	109
precário	PRECARIOUS	C	110
psicologia	PSYCHOLOGY	C	115
psicose	PSYCHOSIS	C	116
puma	PUMA	C	117
quantidade	QUANTITY	C	118
rádio	RADIO	C	119
razão	REASON	C	121
real	REAL	C	122
relacionamento	RELATIONSHIP	C	124
remédio	REMEDY	C	125
requerimento	REQUIREMENT	C	127
rima	RHYME	C	130
robô	ROBOT	C	132
rude	RUDE	C	134
rumor	RUMOR	C	135
samba	SAMBA	C	136
sentimento	SENTIMENT	C	137
significância	SIGNIFICANCE	C	139
similar	SIMILAR	C	140
simples	SIMPLE	C	141
temperatura	TEMPERATURE	C	144
tomate	TOMATO	C	147
tradução	TRANSLATION	C	148
trânsito	TRANSIT	C	150
triunfo	TRIUMPH	C	152
troféu	TROPHY	C	153
veículo	VEHICLE	C	158
vogal	VOWEL	C	163
voto	VOTE	C	164

Appendix G: RStudio `lmer` model raw output

LMER Model

```
install.packages(c("lmerTest", "optimx", "nloptr", "MuMIn", "r2glmm", "car"))

## Installing packages into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.5'
## (as 'lib' is unspecified)

#Loading required packages
library(foreign)
library(lmerTest)

## Loading required package: lme4

## Loading required package: Matrix

##
## Attaching package: 'lmerTest'

## The following object is masked from 'package:lme4':
##
##     lmer

## The following object is masked from 'package:stats':
##
##     step

library(optimx)
library(nloptr)
```

```
library(MuMIn)
library(r2glmm)
library(car)

## Loading required package: carData

#Data optimizer
defaultControl <- list(algorithm="NLOPT_LN_BOBYQA",xtol_rel=1e-6,maxeval=1e5)

nloptwrap2 <- function(fn,par,lower,upper,control=list(),...) {
    for (n in names(defaultControl))
      if (is.null(control[[n]])) control[[n]] <- defaultControl[[n]]
    res <- nloptr(x0=par,eval_f=fn,lb=lower,ub=upper,opts=control,...)
    with(res,list(par=solution,
                  fval=objective,
                  feval=iterations,
                  conv=if (status>0) 0 else status,
                  message=message))
}

#Reading the data file
setwd("/cloud/project/LMER Model")
Dissertation_Dataset <- read.csv("/cloud/project/Master Datasets/Dissertation_Master_Dataset.csv", sep = ",
", dec = ".", encoding = "UTF-8")

#Configuring the data variables
Dissertation_Dataset$RT <- as.numeric(Dissertation_Dataset$RT)
Dissertation_Dataset$invRT <- as.numeric(Dissertation_Dataset$invRT)
Dissertation_Dataset$PAIR <- as.factor(Dissertation_Dataset$PAIR)
Dissertation_Dataset$LIST <- as.factor(Dissertation_Dataset$LIST)
```

```
Dissertation_Dataset$ZTime <- as.numeric(Dissertation_Dataset$ZTime)
Dissertation_Dataset$Time <- as.numeric(Dissertation_Dataset$Time)
Dissertation_Dataset$AGE <- as.numeric(Dissertation_Dataset$AGE)
Dissertation_Dataset$GENDER <- as.factor(Dissertation_Dataset$GENDER)
Dissertation_Dataset$SEMANTIC <- as.factor(Dissertation_Dataset$SEMANTIC)
Dissertation_Dataset$SEMANTIC_OLD <- as.factor(Dissertation_Dataset$SEMANTIC_OLD)
Dissertation_Dataset$SEMANTIC_ORIGINAL <- as.factor(Dissertation_Dataset$SEMANTIC_ORIGINAL)
Dissertation_Dataset$ZPHONETIC <- as.numeric(Dissertation_Dataset$ZPHONETIC)
Dissertation_Dataset$PHONETIC_RAW <- as.numeric(Dissertation_Dataset$PHONETIC_RAW)
Dissertation_Dataset$PARTICIPANT <- as.factor(Dissertation_Dataset$PARTICIPANT)
Dissertation_Dataset$TARGET <- as.factor(Dissertation_Dataset$TARGET)
Dissertation_Dataset$PRIME <- as.factor(Dissertation_Dataset$PRIME)
Dissertation_Dataset$ZEnglish_Freq <- as.numeric(Dissertation_Dataset$ZEnglish_Freq)
Dissertation_Dataset$English_Freq <- as.numeric(Dissertation_Dataset$English_Freq)
Dissertation_Dataset$ZPortuguese_Freq <- as.numeric(Dissertation_Dataset$ZPortuguese_Freq)
Dissertation_Dataset$Portuguese_Freq <- as.numeric(Dissertation_Dataset$Portuguese_Freq)
Dissertation_Dataset$logPortuguese_Freq <- as.numeric(Dissertation_Dataset$logPortuguese_Freq)
Dissertation_Dataset$logEnglish_Freq <- as.numeric(Dissertation_Dataset$logEnglish_Freq)
Dissertation_Dataset$ZPROFICIENCY <- as.numeric(Dissertation_Dataset$ZPROFICIENCY)
Dissertation_Dataset$SEMAC <- as.numeric(Dissertation_Dataset$SEMAC)
Dissertation_Dataset$ERS <- as.numeric(Dissertation_Dataset$ERS)
Dissertation_Dataset$GS <- as.numeric(Dissertation_Dataset$GS)
Dissertation_Dataset$OS <- as.numeric(Dissertation_Dataset$OS)
Dissertation_Dataset$Lev <- as.numeric(Dissertation_Dataset$Lev)
Dissertation_Dataset$NLD <- as.numeric(Dissertation_Dataset$NLD)
Dissertation_Dataset$ZGS <- as.numeric(Dissertation_Dataset$ZGS)
Dissertation_Dataset$ZOS <- as.numeric(Dissertation_Dataset$ZOS)
Dissertation_Dataset$ZLev <- as.numeric(Dissertation_Dataset$ZLev)
Dissertation_Dataset$ZNLD <- as.numeric(Dissertation_Dataset$ZNLD)
```

```
Dissertation_Dataset$length_prime_raw <- as.numeric(Dissertation_Dataset$length_prime_raw)
Dissertation_Dataset$length_TARGET_raw <- as.numeric(Dissertation_Dataset$length_TARGET_raw)
Dissertation_Dataset$Zlength_prime_raw <- as.numeric(Dissertation_Dataset$Zlength_prime_raw)
Dissertation_Dataset$Zlength_TARGET_raw <- as.numeric(Dissertation_Dataset$Zlength_TARGET_raw)
Dissertation_Dataset$English_Freq_gmc <- as.numeric(Dissertation_Dataset$English_Freq_gmc)
Dissertation_Dataset$Portuguese_Freq_gmc <- as.numeric(Dissertation_Dataset$Portuguese_Freq_gmc)
Dissertation_Dataset$SEMAC_gmc <- as.numeric(Dissertation_Dataset$SEMAC_gmc)
Dissertation_Dataset$English_Freq_log_gmc <- as.numeric(Dissertation_Dataset$English_Freq_log_gmc)
Dissertation_Dataset$Portuguese_Freq_log_gmc <- as.numeric(Dissertation_Dataset$Portuguese_Freq_log_gmc)
Dissertation_Dataset$length_TARGET_raw_log_gmc <- as.numeric(Dissertation_Dataset$length_TARGET_raw_log_gmc
)
Dissertation_Dataset$Englisg_Freq_gmc <- as.numeric(Dissertation_Dataset$English_Freq_gmc)
Dissertation_Dataset$NLD_gmc <- as.numeric(Dissertation_Dataset$NLD_gmc)
Dissertation_Dataset$NLD_log <- as.numeric(Dissertation_Dataset$NLD_log)
Dissertation_Dataset$NLD_log_gmc <- as.numeric(Dissertation_Dataset$NLD_log_gmc)
Dissertation_Dataset$PHONETIC_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_gmc)
Dissertation_Dataset$PHONETIC_RAW_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_RAW_gmc)
Dissertation_Dataset$PHONETIC_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_gmc)
Dissertation_Dataset$PHONETIC_log <- as.numeric(Dissertation_Dataset$PHONETIC_log)
Dissertation_Dataset$PHONETIC_RAW_log_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_RAW_log_gmc)
Dissertation_Dataset$AGE_gmc <- as.numeric(Dissertation_Dataset$AGE_gmc)
Dissertation_Dataset$Time_ms_gmc <- as.numeric(Dissertation_Dataset$Time_ms_gmc)
Dissertation_Dataset$ERS_gmc <- as.numeric(Dissertation_Dataset$ERS_gmc)
```

```
#Releveling Pertinent Variables
Dissertation_Dataset$SEMANTIC<-relevel(Dissertation_Dataset$SEMANTIC,"CTRL")
Dissertation_Dataset$LANGUAGE<-relevel(Dissertation_Dataset$LANGUAGE,"Functional Monolinguals")

# Running best fit model
LMER_best_fit_model <- lmer(invRT ~ LANGUAGE*SEMANTIC + ZNLD_gmc*ZPHONETIC_gmc + ZEnglish_Freq_gmc + Zlengt
h_TARGET_raw_gmc + ZTime_ms_gmc + OTHER_LANGUAGES + (1|PAIR) + (1|PRIME) + (1|TARGET) + (SEMANTIC|PARTICIPA
NT) + (0+ZEnglish_Freq_gmc|PARTICIPANT) + (0+ZTime_ms_gmc|PARTICIPANT) + (0+Zlength_TARGET_raw_gmc|PARTICIP
ANT), data = Dissertation_Dataset, REML = FALSE, control=lmerControl(optimizer="nloptwrap2"))

#Obtaining model summary
summary(LMER_best_fit_model)

## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula:
## invRT ~ LANGUAGE * SEMANTIC + ZNLD_gmc * ZPHONETIC_gmc + ZEnglish_Freq_gmc +
##     Zlength_TARGET_raw_gmc + ZTime_ms_gmc + OTHER_LANGUAGES +
##     (1 | PAIR) + (1 | PRIME) + (1 | TARGET) + (SEMANTIC | PARTICIPANT) +
##     (0 + ZEnglish_Freq_gmc | PARTICIPANT) + (0 + ZTime_ms_gmc |
##     PARTICIPANT) + (0 + Zlength_TARGET_raw_gmc | PARTICIPANT)
##    Data: Dissertation_Dataset
## Control: lmerControl(optimizer = "nloptwrap2")
##
##      AIC      BIC   logLik deviance df.resid
##  14364.2  14633.7  -7150.1  14300.2    33484
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -7.4402 -0.6202 -0.0386  0.5881  6.1591
##
## Random effects:
##  Groups        Name                   Variance  Std.Dev. Corr
##  PRIME         (Intercept)            0.0043107 0.06566
##  PAIR          (Intercept)            0.0022357 0.04728
##  TARGET        (Intercept)            0.0092168 0.09600
##  PARTICIPANT   Zlength_TARGET_raw_gmc 0.0004832 0.02198
##  PARTICIPANT.1 ZTime_ms_gmc           0.0152537 0.12351
##  PARTICIPANT.2 ZEnglish_Freq_gmc      0.0001005 0.01003
##  PARTICIPANT.3 (Intercept)            0.0655086 0.25595
##                SEMANTICCOG            0.0031251 0.05590   0.46
##                SEMANTICDISTR          0.0170481 0.13057  -0.36  0.19
##                SEMANTICFF             0.0010710 0.03273   0.54  0.63  0.22
##  Residual                             0.0828356 0.28781
## Number of obs: 33516, groups:
## PRIME, 901; PAIR, 852; TARGET, 591; PARTICIPANT, 57
##
## Fixed effects:
##                                 Estimate Std. Error         df t value
## (Intercept)                   -1.713e+00  6.388e-02  7.091e+01 -26.818
## LANGUAGEBilinguals             1.986e-01  7.636e-02  6.402e+01   2.601
## SEMANTICCOG                   -6.777e-02  1.964e-02  2.878e+02  -3.451
## SEMANTICDISTR                  2.437e-01  2.709e-02  7.954e+01   8.997
## SEMANTICFF                    -4.078e-02  1.995e-02  4.420e+02  -2.044
## ZNLD_gmc                      -8.221e-03  6.348e-03  1.987e+03  -1.295
## ZPHONETIC_gmc                  3.245e-03  2.560e-03  5.485e+03   1.267
```

```
## ZEnglish_Freq_gmc                 -1.888e-02  5.354e-03  4.617e+02  -3.527
## Zlength_TARGET_raw_gmc             3.036e-02  5.928e-03  3.667e+02   5.121
## ZTime_ms_gmc                      -5.692e-02  1.679e-02  5.940e+01  -3.389
## OTHER_LANGUAGESYES                 1.226e-01  6.715e-02  5.691e+01   1.826
## LANGUAGEBilinguals:SEMANTICCOG   -4.164e-02  1.789e-02  5.520e+01  -2.328
## LANGUAGEBilinguals:SEMANTICDISTR -4.531e-03  3.546e-02  5.656e+01  -0.128
## LANGUAGEBilinguals:SEMANTICFF     1.761e-02  1.522e-02  5.307e+01   1.157
## ZNLD_gmc:ZPHONETIC_gmc            7.533e-03  3.277e-03  1.671e+03   2.299
##                                  Pr(>|t|)
## (Intercept)                       < 2e-16 ***
## LANGUAGEBilinguals               0.011542 *
## SEMANTICCOG                      0.000643 ***
## SEMANTICDISTR                    9.31e-14 ***
## SEMANTICFF                       0.041527 *
## ZNLD_gmc                         0.195404
## ZPHONETIC_gmc                    0.205051
## ZEnglish_Freq_gmc                0.000462 ***
## Zlength_TARGET_raw_gmc           4.92e-07 ***
## ZTime_ms_gmc                     0.001250 **
## OTHER_LANGUAGESYES               0.073059 .
## LANGUAGEBilinguals:SEMANTICCOG   0.023631 *
## LANGUAGEBilinguals:SEMANTICDISTR 0.898777
## LANGUAGEBilinguals:SEMANTICFF    0.252305
## ZNLD_gmc:ZPHONETIC_gmc           0.021632 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Correlation matrix not shown by default, as p = 15 > 12.
```

```
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

#Obtaining confidence intervals
lower <- coef(summary(LMER_best_fit_model))[,1] + qnorm(.025)*coef(summary(LMER_best_fit_model))[,2]
upper <- coef(summary(LMER_best_fit_model))[,1] + qnorm(.975)*coef(summary(LMER_best_fit_model))[,2]
cbind(coef(summary(LMER_best_fit_model)), lower, upper)

##                                     Estimate  Std. Error         df
## (Intercept)                     -1.713136770 0.063879946   70.91278
## LANGUAGEBilinguals               0.198577234 0.076356564   64.02065
## SEMANTICCOG                     -0.067771841 0.019639952  287.84904
## SEMANTICDISTR                    0.243740424 0.027090214   79.54454
## SEMANTICFF                      -0.040780256 0.019949363  441.95256
## ZNLD_gmc                        -0.008221424 0.006347647 1986.76272
## ZPHONETIC_gmc                    0.003245229 0.002560453 5485.08733
## ZEnglish_Freq_gmc               -0.018883343 0.005354073  461.67348
## Zlength_TARGET_raw_gmc           0.030355769 0.005927736  366.73161
## ZTime_ms_gmc                    -0.056920741 0.016794282   59.40272
## OTHER_LANGUAGESYES               0.122629148 0.067147924   56.90869
## LANGUAGEBilinguals:SEMANTICCOG   -0.041641526 0.017890928   55.20199
## LANGUAGEBilinguals:SEMANTICDISTR -0.004531108 0.035460415   56.55559
## LANGUAGEBilinguals:SEMANTICFF     0.017614627 0.015219558   53.06934
## ZNLD_gmc:ZPHONETIC_gmc            0.007533500 0.003276959 1670.90311
##                                      t value     Pr(>|t|)        lower
## (Intercept)                     -26.8180685 7.466639e-39 -1.838339164
## LANGUAGEBilinguals                2.6006570 1.154209e-02  0.048921118
## SEMANTICCOG                      -3.4507132 6.428742e-04 -0.106265439
## SEMANTICDISTR                     8.9973607 9.310208e-14  0.190644581
```

```
## SEMANTICFF                          -2.0441884 4.152708e-02 -0.079880288
## ZNLD_gmc                            -1.2951924 1.954043e-01 -0.020662583
## ZPHONETIC_gmc                        1.2674434 2.050506e-01 -0.001773166
## ZEnglish_Freq_gmc                   -3.5269117 4.624682e-04 -0.029377133
## Zlength_TARGET_raw_gmc               5.1209716 4.920232e-07  0.018737619
## ZTime_ms_gmc                        -3.3892929 1.249909e-03 -0.089836929
## OTHER_LANGUAGESYES                   1.8262538 7.305927e-02 -0.008978365
## LANGUAGEBilinguals:SEMANTICCOG      -2.3275219 2.363102e-02 -0.076707101
## LANGUAGEBilinguals:SEMANTICDISTR    -0.1277793 8.987769e-01 -0.074032243
## LANGUAGEBilinguals:SEMANTICFF        1.1573678 2.523049e-01 -0.012215159
## ZNLD_gmc:ZPHONETIC_gmc               2.2989299 2.163180e-02  0.001110778
##                                           upper
## (Intercept)                        -1.587934377
## LANGUAGEBilinguals                  0.348233350
## SEMANTICCOG                        -0.029278243
## SEMANTICDISTR                       0.296836268
## SEMANTICFF                         -0.001680224
## ZNLD_gmc                            0.004219735
## ZPHONETIC_gmc                       0.008263624
## ZEnglish_Freq_gmc                  -0.008389553
## Zlength_TARGET_raw_gmc              0.041973918
## ZTime_ms_gmc                       -0.024004553
## OTHER_LANGUAGESYES                  0.254236661
## LANGUAGEBilinguals:SEMANTICCOG     -0.006575952
## LANGUAGEBilinguals:SEMANTICDISTR    0.064970027
## LANGUAGEBilinguals:SEMANTICFF       0.047444413
## ZNLD_gmc:ZPHONETIC_gmc              0.013956222
```

```
#Obtaining R2 values
r.squaredGLMM(LMER_best_fit_model)

## Warning: 'r.squaredGLMM' now calculates a revised statistic. See the help page.

##             R2m      R2c
## [1,] 0.1614303 0.614307

#Obtaining VIF values
vif(LMER_best_fit_model)

##                               GVIF Df GVIF^(1/(2*Df))
## LANGUAGE                  1.940343  1        1.392962
## SEMANTIC                  7.556333  3        1.400829
## ZNLD_gmc                  2.780622  1        1.667520
## ZPHONETIC_gmc             1.540855  1        1.241312
## ZEnglish_Freq_gmc         1.050145  1        1.024766
## Zlength_TARGET_raw_gmc    1.034999  1        1.017349
## ZTime_ms_gmc              1.000053  1        1.000027
## OTHER_LANGUAGES           1.395945  1        1.181501
## LANGUAGE:SEMANTIC         4.502283  3        1.285007
## ZNLD_gmc:ZPHONETIC_gmc    1.519332  1        1.232612
```

```
1/vif(LMER_best_fit_model)

##                               GVIF        Df GVIF^(1/(2*Df))
## LANGUAGE               0.5153727 1.0000000       0.7178946
## SEMANTIC               0.1323393 0.3333333       0.7138629
## ZNLD_gmc               0.3596318 1.0000000       0.5996931
## ZPHONETIC_gmc          0.6489905 1.0000000       0.8055995
## ZEnglish_Freq_gmc      0.9522495 1.0000000       0.9758327
## Zlength_TARGET_raw_gmc 0.9661841 1.0000000       0.9829466
## ZTime_ms_gmc           0.9999467 1.0000000       0.9999734
## OTHER_LANGUAGES        0.7163604 1.0000000       0.8463807
## LANGUAGE:SEMANTIC      0.2221095 0.3333333       0.7782059
## ZNLD_gmc:ZPHONETIC_gmc 0.6581842 1.0000000       0.8112855
```

Appendix H: RStudio `glmer` model raw output

GLMER Model

```
install.packages(c("lmerTest", "optimx", "nloptr", "MuMIn", "r2glmm", "car"))

## Installing packages into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)

#Loading required packages
library(foreign)
library(lmerTest)

## Loading required package: lme4

## Loading required package: Matrix

##
## Attaching package: 'lmerTest'

## The following object is masked from 'package:lme4':
##
##     lmer

## The following object is masked from 'package:stats':
##
##     step
```

```
library(optimx)
library(nloptr)
library(MuMIn)
library(r2glmm)
library(car)

## Loading required package: carData

## Registered S3 methods overwritten by 'car':
##   method                          from
##   influence.merMod                lme4
##   cooks.distance.influence.merMod lme4
##   dfbeta.influence.merMod         lme4
##   dfbetas.influence.merMod        lme4

#Data optimizer
defaultControl <- list(algorithm="NLOPT_LN_BOBYQA",xtol_rel=1e-6,maxeval=1e5)

nloptwrap2 <- function(fn,par,lower,upper,control=list(),...) {
    for (n in names(defaultControl))
      if (is.null(control[[n]])) control[[n]] <- defaultControl[[n]]
    res <- nloptr(x0=par,eval_f=fn,lb=lower,ub=upper,opts=control,...)
    with(res,list(par=solution,
                  fval=objective,
                  feval=iterations,
                  conv=if (status>0) 0 else status,
                  message=message))
}

#Reading the data file
setwd("/cloud/project/GLMER Model")
```

```
Dissertation_Dataset <- read.csv("/cloud/project/Master Datasets/Dissertation_Master_Dataset_GLMER.csv", se
p = ",", dec = ".", encoding = "UTF-8")

#Configuring the data variables
Dissertation_Dataset$RT <- as.numeric(Dissertation_Dataset$RT)
Dissertation_Dataset$invRT <- as.numeric(Dissertation_Dataset$invRT)
Dissertation_Dataset$PAIR <- as.factor(Dissertation_Dataset$PAIR)
Dissertation_Dataset$LIST <- as.factor(Dissertation_Dataset$LIST)
Dissertation_Dataset$ZTime <- as.numeric(Dissertation_Dataset$ZTime)
Dissertation_Dataset$Time <- as.numeric(Dissertation_Dataset$Time)
Dissertation_Dataset$AGE <- as.numeric(Dissertation_Dataset$AGE)
Dissertation_Dataset$GENDER <- as.factor(Dissertation_Dataset$GENDER)
Dissertation_Dataset$SEMANTIC <- as.factor(Dissertation_Dataset$SEMANTIC)
Dissertation_Dataset$SEMANTIC_OLD <- as.factor(Dissertation_Dataset$SEMANTIC_OLD)
Dissertation_Dataset$SEMANTIC_ORIGINAL <- as.factor(Dissertation_Dataset$SEMANTIC_ORIGINAL)
Dissertation_Dataset$ZPHONETIC <- as.numeric(Dissertation_Dataset$ZPHONETIC)
Dissertation_Dataset$PHONETIC_RAW <- as.numeric(Dissertation_Dataset$PHONETIC_RAW)
Dissertation_Dataset$PARTICIPANT <- as.factor(Dissertation_Dataset$PARTICIPANT)
Dissertation_Dataset$TARGET <- as.factor(Dissertation_Dataset$TARGET)
Dissertation_Dataset$PRIME <- as.factor(Dissertation_Dataset$PRIME)
Dissertation_Dataset$ZEnglish_Freq <- as.numeric(Dissertation_Dataset$ZEnglish_Freq)
Dissertation_Dataset$English_Freq <- as.numeric(Dissertation_Dataset$English_Freq)
Dissertation_Dataset$ZPortuguese_Freq <- as.numeric(Dissertation_Dataset$ZPortuguese_Freq)
Dissertation_Dataset$Portuguese_Freq <- as.numeric(Dissertation_Dataset$Portuguese_Freq)
Dissertation_Dataset$logPortuguese_Freq <- as.numeric(Dissertation_Dataset$logPortuguese_Freq)
Dissertation_Dataset$logEnglish_Freq <- as.numeric(Dissertation_Dataset$logEnglish_Freq)
Dissertation_Dataset$ZPROFICIENCY <- as.numeric(Dissertation_Dataset$ZPROFICIENCY)
Dissertation_Dataset$SEMAC <- as.numeric(Dissertation_Dataset$SEMAC)
Dissertation_Dataset$ERS <- as.numeric(Dissertation_Dataset$ERS)
Dissertation_Dataset$GS <- as.numeric(Dissertation_Dataset$GS)
```

```
Dissertation_Dataset$OS <- as.numeric(Dissertation_Dataset$OS)
Dissertation_Dataset$Lev <- as.numeric(Dissertation_Dataset$Lev)
Dissertation_Dataset$NLD <- as.numeric(Dissertation_Dataset$NLD)
Dissertation_Dataset$ZGS <- as.numeric(Dissertation_Dataset$ZGS)
Dissertation_Dataset$ZOS <- as.numeric(Dissertation_Dataset$ZOS)
Dissertation_Dataset$ZLev <- as.numeric(Dissertation_Dataset$ZLev)
Dissertation_Dataset$ZNLD <- as.numeric(Dissertation_Dataset$ZNLD)
Dissertation_Dataset$length_prime_raw <- as.numeric(Dissertation_Dataset$length_prime_raw)
Dissertation_Dataset$length_TARGET_raw <- as.numeric(Dissertation_Dataset$length_TARGET_raw)
Dissertation_Dataset$Zlength_prime_raw <- as.numeric(Dissertation_Dataset$Zlength_prime_raw)
Dissertation_Dataset$Zlength_TARGET_raw <- as.numeric(Dissertation_Dataset$Zlength_TARGET_raw)
Dissertation_Dataset$English_Freq_gmc <- as.numeric(Dissertation_Dataset$English_Freq_gmc)
Dissertation_Dataset$Portuguese_Freq_gmc <- as.numeric(Dissertation_Dataset$Portuguese_Freq_gmc)
Dissertation_Dataset$SEMAC_gmc <- as.numeric(Dissertation_Dataset$SEMAC_gmc)
Dissertation_Dataset$English_Freq_log_gmc <- as.numeric(Dissertation_Dataset$English_Freq_log_gmc)
Dissertation_Dataset$Portuguese_Freq_log_gmc <- as.numeric(Dissertation_Dataset$Portuguese_Freq_log_gmc)
Dissertation_Dataset$length_TARGET_raw_log_gmc <- as.numeric(Dissertation_Dataset$length_TARGET_raw_log_gmc
)
Dissertation_Dataset$Englisg_Freq_gmc <- as.numeric(Dissertation_Dataset$English_Freq_gmc)
Dissertation_Dataset$NLD_gmc <- as.numeric(Dissertation_Dataset$NLD_gmc)
Dissertation_Dataset$NLD_log <- as.numeric(Dissertation_Dataset$NLD_log)
Dissertation_Dataset$NLD_log_gmc <- as.numeric(Dissertation_Dataset$NLD_log_gmc)
Dissertation_Dataset$PHONETIC_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_gmc)
Dissertation_Dataset$PHONETIC_RAW_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_RAW_gmc)
Dissertation_Dataset$PHONETIC_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_gmc)
Dissertation_Dataset$PHONETIC_log <- as.numeric(Dissertation_Dataset$PHONETIC_log)
Dissertation_Dataset$PHONETIC_RAW_log_gmc <- as.numeric(Dissertation_Dataset$PHONETIC_RAW_log_gmc)
Dissertation_Dataset$AGE_gmc <- as.numeric(Dissertation_Dataset$AGE_gmc)
Dissertation_Dataset$Time_ms_gmc <- as.numeric(Dissertation_Dataset$Time_ms_gmc)
Dissertation_Dataset$ERS_gmc <- as.numeric(Dissertation_Dataset$ERS_gmc)
```

```
#Releveling Pertinent Variables
Dissertation_Dataset$SEMANTIC<-relevel(Dissertation_Dataset$SEMANTIC,"CTRL")
Dissertation_Dataset$LANGUAGE<-relevel(Dissertation_Dataset$LANGUAGE,"Functional Monolinguals")
Dissertation_Dataset$RAW_ACCURACY<-relevel(Dissertation_Dataset$RAW_ACCURACY,"incorrect")

#Running best fit model
GLMER_best_fit_model <- glmer(RAW_ACCURACY ~ LANGUAGE*SEMANTIC + ZEnglish_Freq_gmc + ZTime_ms_gmc + Zlength
_TARGET_raw_gmc + ZSEMAC_gmc + (1| PARTICIPANT) + (1|PRIME) + (1|TARGET) + (1|PAIR), data = Dissertation_Da
taset, family=binomial, control=glmerControl(optimizer="nloptwrap2"))

#Obtaining model summary
summary(GLMER_best_fit_model)

## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## RAW_ACCURACY ~ LANGUAGE * SEMANTIC + ZEnglish_Freq_gmc + ZTime_ms_gmc +
##     Zlength_TARGET_raw_gmc + ZSEMAC_gmc + (1 | PARTICIPANT) +
##     (1 | PRIME) + (1 | TARGET) + (1 | PAIR)
##    Data: Dissertation_Dataset
## Control: glmerControl(optimizer = "nloptwrap2")
##
##      AIC      BIC   logLik deviance df.resid
##  12955.9  13090.6  -6462.0  12923.9    33412
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -19.6469   0.0663   0.1250   0.2319   2.8842
##
```

```
## Random effects:
##  Groups      Name        Variance Std.Dev.
##  PRIME       (Intercept) 0.2766   0.5259
##  PAIR        (Intercept) 0.3275   0.5723
##  TARGET      (Intercept) 1.5102   1.2289
##  PARTICIPANT (Intercept) 0.6580   0.8112
## Number of obs: 33428, groups:
## PRIME, 901; PAIR, 852; TARGET, 591; PARTICIPANT, 57
##
## Fixed effects:
##                                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         4.87710    0.23429  20.817  < 2e-16 ***
## LANGUAGEBilinguals                 -0.61333    0.29588  -2.073 0.038183 *
## SEMANTICCOG                         0.28415    0.23245   1.222 0.221563
## SEMANTICDISTR                      -1.37249    0.19154  -7.166 7.74e-13 ***
## SEMANTICFF                          0.18464    0.26908   0.686 0.492601
## ZEnglish_Freq_gmc                   0.26429    0.08528   3.099 0.001941 **
## ZTime_ms_gmc                        0.17440    0.03395   5.136 2.80e-07 ***
## Zlength_TARGET_raw_gmc              0.22030    0.07301   3.017 0.002550 **
## ZSEMAC_gmc                          0.49578    0.13141   3.773 0.000161 ***
## LANGUAGEBilinguals:SEMANTICCOG      0.53983    0.24560   2.198 0.027950 *
## LANGUAGEBilinguals:SEMANTICDISTR -0.13628    0.15150  -0.900 0.368368
## LANGUAGEBilinguals:SEMANTICFF    -0.34842    0.26453  -1.317 0.187791
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
```

```
Correlation of Fixed Effects:
##                      (Intr) LANGUAGEBl SEMANTICC SEMANTICD SEMANTICF ZEn_F_
## LANGUAGEBln          -0.685
## SEMANTICCOG          -0.321  0.195
## SEMANTICDIS          -0.580  0.247      0.394
## SEMANTICFF           -0.271  0.169      0.220     0.330
## ZEnglsh_Fr_          -0.039 -0.004      0.006     0.174    -0.015
## ZTim_ms_gmc           0.026 -0.007     -0.010    -0.005     0.005    -0.011
## Zl_TARGET__           0.025  0.003     -0.032    -0.011     0.050     0.152
## ZSEMAC_gmc           -0.293  0.492     -0.001    -0.003     0.000     0.001
## LANGUAGEB:SEMANTICC   0.241 -0.259     -0.698    -0.293    -0.204     0.004
## LANGUAGEB:SEMANTICD   0.396 -0.430     -0.384    -0.572    -0.332     0.009
## LANGUAGEB:SEMANTICF   0.219 -0.242     -0.217    -0.268    -0.734     0.006
##                      ZTm_m_ Z_TARG ZSEMAC LANGUAGEB:SEMANTICC
## LANGUAGEBln
## SEMANTICCOG
## SEMANTICDIS
## SEMANTICFF
## ZEnglsh_Fr_
## ZTim_ms_gmc
## Zl_TARGET__           0.002
## ZSEMAC_gmc            0.025 -0.002
## LANGUAGEB:SEMANTICC   0.005  0.002  0.002
## LANGUAGEB:SEMANTICD  -0.002  0.008 -0.002  0.511
## LANGUAGEB:SEMANTICF  -0.001  0.003  0.000  0.291
##                      LANGUAGEB:SEMANTICD
## LANGUAGEBln
## SEMANTICCOG
```

```
## SEMANTICDIS
## SEMANTICFF
## ZEnglsh_Fr_
## ZTim_ms_gmc
## Zl_TARGET__
## ZSEMAC_gmc
## LANGUAGEB:SEMANTICC
## LANGUAGEB:SEMANTICD
## LANGUAGEB:SEMANTICF  0.474

#Obtaining confidence Intervals
lower <- coef(summary(GLMER_best_fit_model))[,1] + qnorm(.025)*coef(summary(GLMER_best_fit_model))[,2]
upper <- coef(summary(GLMER_best_fit_model))[,1] + qnorm(.975)*coef(summary(GLMER_best_fit_model))[,2]
cbind(coef(summary(GLMER_best_fit_model)), lower, upper)

##                                    Estimate Std. Error    z value     Pr(>|z|)
## (Intercept)                       4.8770982 0.23428968 20.8165304 3.065976e-96
## LANGUAGEBilinguals               -0.6133302 0.29588170 -2.0728900 3.818252e-02
## SEMANTICCOG                       0.2841483 0.23245440  1.2223829 2.215629e-01
## SEMANTICDISTR                    -1.3724860 0.19153675 -7.1656537 7.741623e-13
## SEMANTICFF                        0.1846397 0.26908458  0.6861772 4.926014e-01
## ZEnglish_Freq_gmc                 0.2642935 0.08527900  3.0991633 1.940680e-03
## ZTime_ms_gmc                      0.1744004 0.03395427  5.1363335 2.801504e-07
## Zlength_TARGET_raw_gmc            0.2203011 0.07301016  3.0174025 2.549510e-03
## ZSEMAC_gmc                        0.4957788 0.13141050  3.7727487 1.614589e-04
## LANGUAGEBilinguals:SEMANTICCOG    0.5398262 0.24560000  2.1979895 2.794985e-02
## LANGUAGEBilinguals:SEMANTICDISTR -0.1362771 0.15149746 -0.8995339 3.683684e-01
## LANGUAGEBilinguals:SEMANTICFF    -0.3484214 0.26452814 -1.3171431 1.877907e-01
##                                        lower       upper
```

```
## (Intercept)                        4.41789884  5.33629750
## LANGUAGEBilinguals                -1.19324768 -0.03341275
## SEMANTICCOG                       -0.17145397  0.73975055
## SEMANTICDISTR                     -1.74789119 -0.99708091
## SEMANTICFF                        -0.34275640  0.71203579
## ZEnglish_Freq_gmc                  0.09714978  0.43143730
## ZTime_ms_gmc                       0.10785130  0.24094959
## Zlength_TARGET_raw_gmc             0.07720376  0.36339835
## ZSEMAC_gmc                         0.23821894  0.75333862
## LANGUAGEBilinguals:SEMANTICCOG     0.05845907  1.02119339
## LANGUAGEBilinguals:SEMANTICDISTR -0.43320667  0.16065247
## LANGUAGEBilinguals:SEMANTICFF    -0.86688703  0.17004422

#Obtaining R2
r.squaredGLMM(GLMER_best_fit_model)

## Warning: 'r.squaredGLMM' now calculates a revised statistic. See the help page.

## Warning: The null model is correct only if all variables used by the original
## model remain unchanged.

##                   R2m       R2c
## theoretical 0.1938626 0.5625173
## delta       0.0744979 0.2161653
```

```
#Obtaining VIF values
vif(GLMER_best_fit_model)

##                              GVIF Df GVIF^(1/(2*Df))
## LANGUAGE                 1.756417  1        1.325299
## SEMANTIC                 5.929369  3        1.345348
## ZEnglish_Freq_gmc        1.087451  1        1.042809
## ZTime_ms_gmc             1.001749  1        1.000874
## Zlength_TARGET_raw_gmc 1.036538  1        1.018105
## ZSEMAC_gmc               1.425754  1        1.194050
## LANGUAGE:SEMANTIC        6.720202  3        1.373716


1/vif(GLMER_best_fit_model)

##                               GVIF        Df GVIF^(1/(2*Df))
## LANGUAGE                 0.5693409 1.0000000       0.7545468
## SEMANTIC                 0.1686520 0.3333333       0.7433019
## ZEnglish_Freq_gmc        0.9195813 1.0000000       0.9589480
## ZTime_ms_gmc             0.9982537 1.0000000       0.9991265
## Zlength_TARGET_raw_gmc 0.9647503 1.0000000       0.9822170
## ZSEMAC_gmc               0.7013831 1.0000000       0.8374862
## LANGUAGE:SEMANTIC        0.1488050 0.3333333       0.7279523
```