Toes in Pies: Exploring the Processing of Spanish – English Interlingual Homographs

Nikki Dean Marshall

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

February 2021

CANDIDATE DECLARATION

I hereby declare that:

1. I have not been enrolled for another award of the University, or other academic or professional organisation, whilst undertaking my research degree.
2. None of the material contained in the thesis has been used in any other submission for an academic award.
3. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.
2. The work undertaken towards the thesis has been conducted in accordance with the SHU Principles of Integrity in Research and the SHU Research Ethics Policy.
3. The word count of the thesis is 64,747

Name	*Nikki Dean Marshall*
Date	*February 2021*
Award	*PhD*
Faculty	*Social Sciences and Humanities*
Director(s) of Studies	*Dr Anna Maria Di Betta*

Abstract

The interplay of bilinguals' languages during visual written word recognition was investigated using a novel set of Spanish (L1) – English (L2) interlingual homographs, which are words ambiguous cross-linguistically. Differently from previous literature, both identical (e.g., PIE, meaning *foot* in Spanish) and near-identical (e.g., CARPET, with CARPETA meaning *folder* in Spanish) interlingual homographs were considered and a database was created to allow for an easy-accessible list of stimuli including their associated linguistic characteristics in both languages. A series of studies were conducted in the participants' L2 at both the lexical and the semantic level to explore the extent of the non-target L1's activation and the role of cross-linguistic orthographic overlap. Study 1's findings showed that L1 linguistic properties significantly predicted bilinguals' performance in lexical decision task. In Study 2 interlingual homographs were used as primes and inserted at the end of sentences biased to the L2 reading; and were followed by targets related to either the Spanish, English, or unrelated to either meaning. Prime duration was manipulated to explore late (500ms; Study 2a) and early (200ms; Study 2b) stages of processing. Bilinguals showed significant negative priming for the Spanish-related meaning in the 500ms prime condition but not in the 200ms. Overall, Study 1 and 2 found no significant processing differences between identical and near-identical homographs.

Study 3, instead, investigated within-language ambiguity by using homonyms, which, like interlingual homographs, have the same word form but different meanings. Like Study 2, homonyms were placed at the end of sentences biased to the non-dominant meaning and stimuli were created to mimic near-identical interlingual homographs. Native English speakers produced significant priming for the sentence-relevant meaning only, irrespective of whether meaning activation was probed at early or later processing stages, thus suggesting a swift resolution of ambiguity in L1.

Findings are discussed with reference to existing models of the bilingual lexicon (e.g., BIA+ model, Dijkstra & van Heuven, 2002). In particular, evidence of the activation of the non-target L1 was found (i.e., Studies 1b and 2a) which supports the idea of non-selective access.

I dedicate this doctoral thesis to my loving and supportive family, in particular, to my Mum and Dad, and Grandma Jean Marshall.

Acknowledgements

There are many wonderful people whom I would like to thank for their contributions, both directly and indirectly, to this doctoral thesis, to only some of whom it is possible to give particular mention here.

Above all, I would like to thank my parents, for their unequivocal support throughout my life and academic studies. Mum, thank you for all the calls and messages, motivational journals, and for just being you: the kindest, most supportive and fabulous woman I know. Will never forget your whiteboard “to do” list parades; your excitement about a completed task tick, you have lightened grey times with much needed humour. You have always pushed me to be the best I can be, I cannot thank you enough for never giving up on me. Dad, thank you for always pretending to know what I am jabbering on about, and proudly supporting your Scoot. Your gardening tip messages, and funny jokes were just the ticket to bring a smile to my face at the end of many a stressful day. I look forward to celebrating with you in the Lemon Tree!

This thesis would not have been possible without my academic family. I am honoured to be able to thank my Director of Studies, Dr Anna Maria Di Betta, and my second supervisor, Dr Jane Morgan, for their outstanding guidance, emotional and academic support, and endless patience. Throughout this journey, you have both been my biggest cheerleaders and critics, challenging me to improve at every step but with the same objective, for which we are finally here. Your voices will forever live on in my writing. Thank you for being the dream team, and for rekindling my love of research through your own passion; you are both truly exceptionally inspirational people. I would also like to thank Dr David Playfoot for his friendship and sound sensible advice which has been invaluable on both an academic and a personal level; I will be eternally grateful. You once said that the beauty of the stars is lost without the darkness in between. Thank you for being my calm in the darkness. Overall, I could not have asked for a better team to intellectually challenge but also to nurture and guide me through this experience. Thank you from the bottom of my heart. I very much hope that we will have the opportunity to work together in the future.

During my doctoral thesis journey, I have been lucky to live with some amazing characters. Rachel, thank you for the years of friendship, may there be many more. To the tripod, Oli and Rosie, thank you for always triggering laughter, even at the most stressful of times (e.g., brie and cranberry sandwich saga!). Jersey, thank you for being my rock in the last year. While my rescue dog Rudie may be unable to read this, he has, through his required comfort and exercise breaks, given me my own time away to enable me to collect my thoughts ready for the next chapter of writing. I would like to thank him and Jersey for keeping me sane during the last months of my writing.

Thank you to all my friends and colleagues that have supported me. A special mention to the PSP PhD Community, I feel lucky to be part of such a supportive group – you are all “SHINE” award worthy! Thank you to Percy the PSP mascot, a little dog with an immense amount of character, who I miss greatly. I would also like to give a special thanks to a couple of people. To "Chuckle Sister" Lucy, for the climbs, Fredos, and chats. Fabulous Ben, for the talks and fun research ideas we are now embarking on. To Lonsdale ($N^2 > NR^2$), thank you for the lucky desk, the post-it notes, and the entertaining conversations. To Becca W, for your support and being the best “neighbourino”. To Anna R, for your infinite enthusiasm and encouragement. To Fran, for never asking how the PhD was going. To Holly, for being a breath of fresh air when needed. To Roxy, for always listening with Fanta Limón and Reggaetón. To Rosie M, for being my lobster. Beth, you once gave me a frame from our memorable walk with Alex and attached a quote “Whenever you’re feeling stressed, remember if you can survive this walk, you can survive the PhD”. That picture has been on my desk since, and how right you were Beth, thank you. Thank you to the fellow bonfire loving Ledsam, for your entertaining and intellectual conversations, and joining me in the fire dancers – may our future flames be smaller but never dull. I would also like to thank Pseudohomophone girl, your endless support and proofreading got me to the finish line. I am pleased that you can now tell Papa and Mama Porter that it is done. I look forward to offering you the same support in your doctoral journey. To my dear Sam, thank you for your unwavering friendship and being a constant through the years.

I would like to acknowledge the demonstratorship fund that helped me financially and gave me invaluable teaching experience. Additionally, thank you to the Doctoral School and the PSP department; your funding helped me go to European and International conferences that I could otherwise only have dreamt of. Thank you for giving me the opportunity to represent our university and disseminate my research. I would also like to thank the PGRT team for their help, specifically, Dr Lynne Barker for her advice throughout the years. I will not forget the technical resources team, Lee and Jordan (and Dan) for their legendary “window chats”. Always ready to listen and help, you really are the heart of the campus, thank you.

Me gustaría agradecer a mis examinadores, Dr Cristina Izura y Dr Javier Aguado-Orea, por tomarse el tiempo para leer y examinar mi tesis doctoral con tan interés. Gracias por tus pensamientos y discusiones en la viva, aprendí mucho y disfruté hablando con ustedes. I would also like to thank my viva Chair Professor Lisa Hopkins for taking the time to oversee my online viva. Last but not least, I would like to thank my participants. So many great people have supported me through this doctoral thesis but without you, there would be no data and only theory. Thank you for taking part and for your interesting discussions – you are the heart of this research.

Table of contents

Table of Tables

Table of Figures

Chapter 1 – General Introduction.

Humans are surrounded by language in nearly every aspect of their lives and use it to communicate their thoughts, feelings, and to understand the world around them (Marian & Shook, 2012). For many people, this rich linguistic environment is not limited to just one language; it can involve up to two or more languages. It is estimated that as much as two-thirds of the people in the world are bilingual (Shin, 2017), and this number is progressively rising (Eurobarometer, 2006; Grosjean, 2010) suggesting that bilinguals are the norm compared to monolinguals. Bilingualism research has found that bilinguals have an advantage compared to monolinguals in the domain of cognitive control (Blumenfeld & Marian, 2014). Cognitive control is defined as the regulations and coordination of thoughts to respond suitably to stimuli in the environment and to maintain focus on goal directed behaviour (e.g., relating to domains such as working memory, inhibition, etc; Braver, 2012). Bilinguals have to manage two (or more) language systems (Yang, 2017); however interestingly, bilinguals appear to be able to function in each of their languages separately when required to do so, with limited or no apparent interference from the language that is not being used (i.e., non-target language; Green, 1998). Moreover, bilinguals can switch easily from one language to the next language (also known as code-switching, for theoretical reasons why bilinguals code-switch, see Heredia & Altarriba, 2001), and be able to recognise language switches when changing from one language to another (Abutalebi & Green, 2008). Consequently, bilinguals have this extra training in cognitive control skills compared to monolinguals which is argued to contribute to this bilingual advantage in cognitive control (Struys et al., 2018). Therefore, given that bilinguals are more common compared to monolinguals (Shin, 2017) and that there are processing differences between them, it is unsurprising that there is an abundance of interest in bilingualism.

Additionally, researching bilinguals raises questions that are not applicable to individuals with only one language, such as how bilinguals distinguish words from the two languages while reading and speaking.

There are two questions that dominate the field of bilingualism in relation to how languages are stored and accessed in an individual's mental lexicon. The first question centres on storage and refers to the fact that it is unclear whether all words are *stored* in an integrated lexicon, regardless of which language the word belongs to, or whether there are separate lexicons for each known language. The lexicon refers to the representation of all properties of words such as, for instance, meaning and pronunciation in both languages (Jackendoff & Jackendoff, 2002). The second question centres on access and refers to how a bilingual can *select* the correct language to use depending on context, and whether all the languages the bilingual knows become active during lexical retrieval instead of just the language that is relevant.

To answer such questions, researchers have capitalised on words that are shared across languages such as *cognates*, which share form and meaning across languages (such as the word PIANO in English and Spanish), and *interlingual homographs* that, instead, share form but have different meaning across languages (such as PIE, that means *foot* in Spanish but refers to a type of food in English). While cognates are not the stimuli of interest in this research programme, their findings have made an important contribution by informing the debate of how languages are stored and accessed by bilinguals. Moreover, cognates have been researched in more depth in comparison to interlingual homographs, which are the stimuli of interest of this thesis. Briefly, research on cognates has found that they are processed more quickly compared to matched control words (for a review, see Poort & Rodd, 2017) and this is referred to as the *cognate facilitation effect.* This effect provides evidence that both languages are activated and can influence current language processing. The cognate facilitation effect is robust and well documented in a variety of languages and tasks such as visual word recognition (e.g., Cristoffanini, Kirsner, & Milech, 1986; De Groot & Nas, 1991; Dijkstra, Grainger, & van Heuven, 1999; Peeters, Dijkstra, & Grainger, 2013; Voga & Grainger, 2007) and recognition in sentence paradigms (e.g., Duyck, van Assche, Drieghe, & Hartsuiker, 2007; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009; Van Hell & De Groot, 2008).

Moreover, the cognate facilitation effect is found to be consistent in both native language (L1; e.g., Van Hell & Dijkstra, 2002) and second language (L2) processing (e.g., Dijkstra et al., 1999; Lemhöfer & Dijkstra, 2004; Peeters et al., 2013). For instance, priming a cognate in one language speeds up the processing of the other language as well (e.g., Bowers, Mimouni, & Arguin, 2000; Cristoffanini et al., 1986; Gerard & Scarborough, 1989; Poort, Warren, & Rodd, 2016), though the facilitatory priming effect is generally stronger in non-native languages. The orthography of cognates and its effects on the cognate facilitation effect have also been explored, and generally, the overall magnitude of cognate facilitation effects is moderated by differences in form overlap. While both identical and near-identical cognates (the latter being words such as, e.g., the MAPA in Spanish and MAP in English) are processed more quickly compared to matched controls, stronger facilitation effects are produced for identical cognates compared to near-identical cognates (e.g., van Assche, Drieghe, Duyck, Welvaert & Hartsuiker, 2011; Dijkstra, Miwa, Brummelhuis, Sappelli, Baayen, 2010; Peeters et al., 2013; Vanlangendonck, Peeters, Rueschemeyer, & Dijkstra, 2020). Thus, the facilitation effect reduces in magnitude with decreasing form overlap.

The cognate facilitation effect is understood to occur by assuming that both orthographic representations of an identical cognate become activated by the input, and this subsequently converges activation to a shared semantic representation (Dijkstra et al., 2010; Dijkstra et al., 2019, Vanlangendonck et al., 2020). For example, when a Spanish – English bilingual sees the English word MAP it may activate the stored orthographic representation of both MAP and MAPA (the latter being the Spanish translation equivalent), but as both have the same meanings, this speeds up recognition for bilinguals because the more orthography the word forms share, the more strongly they activate at the semantic level. Faster responses compared to control words stems from overlapping semantic representations; overlapping orthographic representations enhance this facilitation effect (Dijkstra et al., 2019), hence it is the spelling moderating the facilitation, words that are less similar in orthography will see a decrease in facilitation as there is not as much overlap between the two representations.

Although cognate research provides insight into bilingual language processing, it is difficult to tease apart the effects of orthography and semantics because they share both elements. In contrast to cognates, interlingual homographs are words that share spelling across languages but not meanings, and therefore can provide an opportunity to separate the effects of orthography and semantics across languages. Researchers have made use of interlingual homographs to study bilingual language processing in a variety of tasks: lexical decision tasks (LDTs; e.g., De Bruijn, Dijkstra, Chwilla, & Schriefers, 2001; De Groot, Delmaar, & Lupker, 2000; Dijkstra, De Bruijn, Schriefers, & Ten Brinke, 2000a; Dijkstra et al., 1999; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Lemhöfer & Dijkstra, 2004; Poort et al., 2016), semantic relatedness judgement tasks (e.g., Macizo, Bajo, & Martín, 2010), and sentence contexts (e.g., Libben & Titone, 2009; Titone, Libben, Mercier, Whitford, & Pivneva, 2011). However, as it will be shown in the following chapters, interlingual homograph findings have been more inconsistent compared to those obtained with cognates which generally find facilitation effects.

Additionally, interlingual homographs that are not identical in spelling (so called near-identical interlingual homographs; for example, the word CARPET, which refers to a type of flooring in English, and the Spanish word CARPETA, which means *folder* in Spanish) have not been explicitly studied. Thus, given that the processing of cognates is altered by the degree of spelling overlap, an open question exists concerning whether near-identical interlingual homographs are processed in the same manner as their identical counterparts.

1.1 Thesis Overview

The current research aims to investigate the relative activation and impact of the L1 on L2 processing using a novel set of Spanish–English identical and near-identical interlingual homographs. The English and Spanish languages were chosen because the researcher is a native speaker of both languages and because of their prevalence compared to other languages worldwide. It is estimated that worldwide there are 1.35 billion English speakers and 543 million Spanish native or L2 speakers (Ethonologue, 2021; Szmigiera, 2021). English is the most spoken language worldwide, and there are only two other languages that are spoken more compared to Spanish: Chinese Mandarin (1.12 billion speakers) and Hindi (600 million speakers; Ethonologue, 2021; Szmigiera, 2021). While Chinese Mandarin and Hindi have more speakers worldwide compared to Spanish, the writing systems and alphabets are different to English. Whereas the Spanish and English languages have similar writing systems as they share Indo-European heritage, this makes it easier to investigate interlingual homographs as the shared alphabet allows for lexical similarity across languages (Colorado, 2007). There are also diverging features between English and Spanish that make this combination of languages interesting to research (for a review of similarities and differences in orthography, see Pérez Cañado, 2016). Although both orthographic systems have alphabetic foundations, the Spanish language is more transparent in its orthography (i.e., the spelling of words mirrors their pronunciation) presenting with more notable phoneme-grapheme correspondence compared to the English language, which lacks orthographic regularity (Pérez Cañado, 2016). English is complex because of this irregularity and requires speakers to have an understanding of morphemic structure and conventions governing representation of the orthography (Barry, 1992; Goulandris, 1992; Pérez Cañado, 2000; Seymour, 1992). Thus, overall, exploring the relationship between an irregular and regular language that also share features makes it an interesting combination to research. Given the lexical similarities between the two languages, one might expect to see increased levels of activation from the non-target language when lexical features are shared, such as in the case of identical interlingual homographs). However, this activation could be reduced in instances where the two languages diverge orthographically, such as in the case of near-identical interlingual homographs.

Interlingual homographs were chosen as stimuli to add cross-linguistic ambiguity at the semantic level, and therefore allow one to explore their processing with either lexical or semantic tasks. In addition, one of the novelties of this research programme is that it makes use of both identical and near-identical interlingual homographs which have not been investigated separately in previous research at a lexical *and* semantic level, apart from limited research using semantic tasks (Di Betta et al., 2015). Furthermore, to mirror the cross-linguistic studies, this research programme also explores language ambiguity *within* a language by using English homonyms, which are words that, like interlingual homographs, have two or more different meanings making them ambiguous (e.g., the word BARK which can refer to the sound produced by dogs and to the outer layer of a tree; Rodd, 2017). Using stimuli from the monolingual domain not only allows parallels to be drawn between different types of ambiguous words but presents the possibility that principles from models of semantic ambiguity can be applied to the bilingual processing (Rodd, 2017). In this research programme this comparison of between and within-language effects has the added strength that the same task paradigm is used (see Figure 1 for a visual conceptualisation of chapter 4 and 5's methodological design) which allows for cross-comparison between ambiguous stimuli.

The remainder of Chapter 1 will discuss the different definitions of bilingualism that have been proposed in the literature, although it will not attempt to provide an exhaustive definition of bilingualism (e.g., Beardsmore, 1986; Hornby, 1977; Skutnabb-Kangas, 1981; Romaine, 1989; Hoffmann, 2014; Baker, 2007). This chapter, instead, aims to offer the reader through the definitions the necessary tools to interpret previous research drawn from bilingual populations, and to provide information about how this research programme will define and operationalise bilingualism (for a short review of the classification of bilingualism, see Moradi, 2014).

Up to now, most of the research looking at how languages are stored and accessed supports the idea that bilinguals store their languages in an integrated lexicon (also known as the shared-lexicon account; e.g., Dijkstra & van Heuven, 2002; Salamoura & Williams, 2007) and that words from both languages become active during lexical retrieval (also known as the non-selective access account; for a review, see Dijkstra, 2005; Dijkstra & van Heuven, 2012; Dijkstra et al., 2019). Chapter 2 aims to provide the reader with a theoretical context. It will therefore evaluate the different models of the bilingual lexicon that have been proposed and identify the framework best suited to interpret the findings of the studies presented in this thesis.

Chapter 3 introduces the set of Spanish-English interlingual homographs that will form the stimuli for the lexical and semantic tasks for this research programme, together with their psycholinguistic characteristics in both languages. Although, because of the novel nature of near-identical interlingual homographs, a pilot study is first conducted to validate the ambiguous nature of these words. As mentioned above, one of the originalities of present work is that *both* identical and near-identical interlingual homographs will be considered as they can lead to potentially different predictions in terms of the interplay between the two languages. Namely, identical interlingual homographs are words that belong to both languages and subsequently are likely to activate both L1 and L2, whilst the near-identical interlingual homographs are only real words in the L2 English, and therefore the activation of the L1 Spanish is likely to be reduced or non-existing. Different from most of the previous research using LDTs, this chapter will explore the activation of the two languages by identifying the L1 and L2 psycholinguistic variables that best account for the bilinguals' performance when making lexical decisions in one language only. Significant involvement of L1 variables will indicate activation of the non-target language during the task at hand which involves the L2. The relationship between orthographic overlap and L1 and L2 linguistic characteristics has not previously been explored in identical and near-identical interlingual homographs; and thus, a by-product of this study is the development of a set of interlingual homographs with associated psycholinguistic variables that is intended to be a useful tool for future research.

Chapter 4 builds on the previous chapter’s work by investigating the processing of interlingual homographs at the semantic level using a written sentence priming and lexical decision paradigm. The aim of this study is to explore whether strong activation of (and therefore interference from) the irrelevant L1 Spanish can be found in a task that requires semantics, and whether the pattern of results is modulated by orthographic overlap. One might expect that in an exclusively L2 context, when a sentence context is constrained and primes the L2 meaning of the interlingual homograph, that there would be minimal or no activation of the L1 meaning. To further explore this hypothesis, the activation of the participants' L1 and L2 will be probed at different points by manipulating the prime duration to investigate early (200 ms) and later (500 ms) language processing.

Chapter 5 moves the focus from the bilingual to the monolingual lexicon and is a conceptual replication of the study reported in chapter 4 (see Figure 1 for a visual conceptualisation of the stimuli and design). Instead of an exact replication where identical method, materials, and stimuli are used again, a conceptual replication attempts to reproduce the methods of an earlier study as closely as possible (Diener & Biswas-Diener, 2016; Shrout & Rodgers, 2018). Therefore, a conceptual replication tests the same idea making a minor change to the method, so if the same or similar results are found they indicate that the original findings are robust to an alternative methodological change, and thus seeing how generalizable the findings truly are (e.g., Diener & Biswas-Diener, 2016; Shrout & Rodgers, 2018; Stroebe, 2016; 2019). Namely, based on the suggestions that interlingual homographs can be compared to homonyms (Rodd, 2017), this chapter explores within-language ambiguity in native English speakers (NEs). It aims to investigate whether activation of the homonym’s different meanings can be probed at different time points using a written sentence priming and LDT. Additionally, it aims to investigate the processing of stimuli developed to mirror near-identical interlingual homographs.

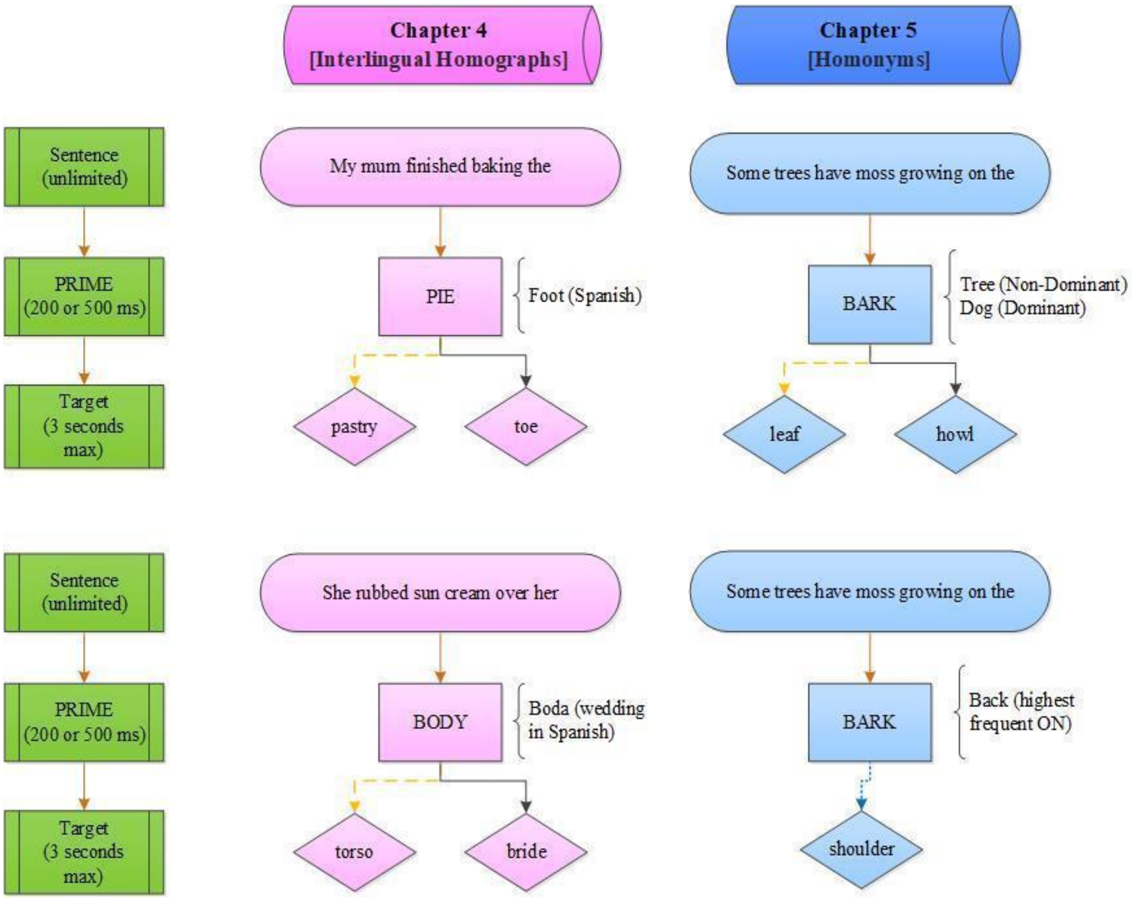

Figure 1.

Visual conceptualisation of the experimental stimuli and design used in Chapters 4-5.

Timings are represented in brackets in Figure 1. In chapter 4, Spanish – English bilinguals read sentences biased to the L2 meaning of the homograph, followed by a homograph prime, and then the target that was either related to the English meaning, Spanish meaning, or unrelated. Whereas, in chapter 5, sentences are biased towards the non-dominant meaning of the homonym and the target is either related to the dominant meaning of the homonym, the non-dominant meaning, related to the ON, or is unrelated to either meaning of the homonym.

Finally, Chapter 6 will outline the concluding remarks and discuss future avenues of research. All the experimental chapters (3 – 5) are designed to be submitted for publication to peer-reviewed journals.

1.2 Characteristics that form the bilingual experience

In its simplest form, bilingualism can be defined as *knowing* two languages (Valdés & Figueora, 1994), though complications arise when defining what it means to *know* a language. It is impossible to find two bilinguals who share identical language experiences and proficiency, making it challenging to identify a single agreed upon definition of bilingualism. The lack of a definition agreement has left researchers disputing what it is that truly encompasses bilingualism in an attempt to generate a consistent description. One of the first attempts was by Bloomfield (1933) who defined bilingualism as native–like control of two languages (also known as balanced bilinguals), while, in contrast, Mackey (1962) defined it as the ability to use more than one language. Similarly, to Mackey (1962), Weinreich (1968) defined bilingualism as the practice of alternately using two languages, whilst Haugen (1953) proposed "the point where a speaker can first produce complete meaningful utterances in the other language" (p.7) to be key for defining bilingualism.

As can be seen, these definitions range from Bloomfield's (1933) rigorous expectations of balanced bilingualism to Mackey's (1962), Weinreich's (1968) and Haugen's (1953) less stringent requirements of basic ability or the practice of using two languages. Moreover, these definitions do not take into account individual differences between bilinguals, and this is one of the reasons the literature classifies bilinguals according to the age when the individuals were first exposed to a L2, distinguishing between early and late bilinguals. It should be noted that the cut-off points to be classified as an early or late bilingual are again not clearly stated in the literature. Therefore, for the purposes of this research programme, the acquisition of more than one language in the pre-adolescent phase of life (i.e., childhood) is defined as early bilingualism (Beardsmore, 1986), whereas the acquisition of an L2 after puberty is defined as late bilingualism (Hoffmann, 2014). In other words, the native language of a late bilingual has already been fully established when they learn their L2.

The early and late bilingual definitions are also associated with individual attainments of linguistic competence. For example, early bilinguals are commonly regarded as attaining native-like linguistic competence in both languages, and have therefore, two native languages (Moradi, 2014). In contrast, late bilinguals are regarded as non-native speakers of the L2, who generally have not obtained complete competence of L2 as they have learned it after the *critical period* of puberty (Hoffmann, 2014; Lenneberg, 1967). This *critical period* where language develops readily is argued to happen between the age of two and at the age of puberty (Hoffmann, 2014; Lenneberg, 1967), and language acquisition is regarded as more difficult and therefore less successful after this period. For example, this means being more prone to grammatical errors and having a stronger foreign accent (e.g., Hartshorne, Tenenbaum, & Pinker, 2018; Robson, 2002).

Exploring the definitions of bilingualism is important because it has influenced the theoretical underpinnings of models that attempt to understand how languages are stored and accessed. For instance, in one of the pioneering studies of bilingualism, Weinreich (1953) proposed a model concerning the organisation of the bilingual memory, which included the conditions which lead to differences in the bilingual structure of the lexicon. These conditions are centred on the environment in which the languages are acquired, because even though bilingual individuals share the overall experience of using two languages in their lives, the ways in which they acquire them may differ (Beardsmore, 1986). Thus, a distinction was made between three types of bilinguals: coordinate, compound, and subordinate (see Figure 2).

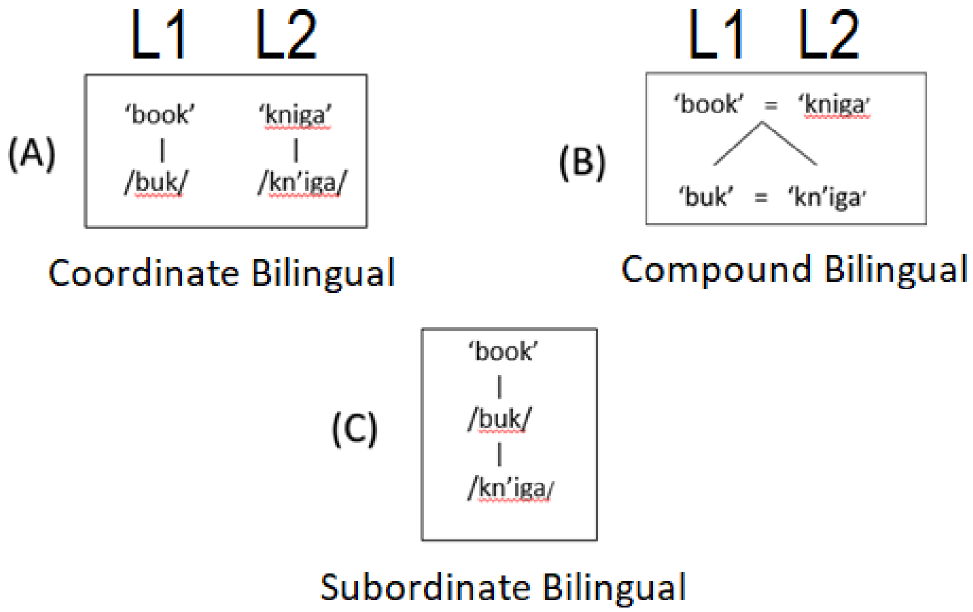

Figure 2.
Weinreichs' (1953) three types of bilinguals: Coordinate (A), Compound (B) and Subordinate (C). Source: Adapted from Kroll and Stewart (1994).

In Figure 2 the examples are from the English and Russian language. Bilingual A: 'book' and 'kniga' are the semantic components of a word and /buk/ and /kn'iga/ refer to the lexical forms in the two languages; this type of bilinguals' stores languages separate in terms of both orthography and semantics. Bilingual B: the lexical component of a word is stored separately but meanings of a word are stored in an integrated fashion. Bilingual C: the meaning of a L2 word is accessed through the corresponding L1 counterpart.

Firstly, with *coordinate bilinguals* (Bilingual A), each language is stored in two meaning units; in other words, one system of meaning for words that the person knows in their L1 and the other for words in their L2. This bilingual structure is sometimes referred to as pure bilingualism because the bilinguals' languages are independent both at the conceptual and lexical levels (Heredia & Cieślicka, 2014). For instance, in the case of a Spanish - English bilingual, the structure would suggest that the definitions of "ordenador" and "computer" represent meanings that may be unique to each language. Therefore, the Spanish word "ordenador" and its English translation "computer" are different, and both meanings are associated with language-specific information. Since each language is learned as an independent communication system, the switch between the two semantic systems is dependent on context because the languages are used for different functions. For example, some children use their L1 exclusively in the family home and their L2 in a school environment (Tabors & Snow, 2001). The learning context is fundamental in determining the bilingual’s situation because the information is retrieved in the language context in which it was encoded (Kolers, 1968). In contrast, while *compound bilinguals* (Bilingual B) have independent systems for information at the lexical or word level, information at the meaning or conceptual level is shared. For instance, for a Spanish - English compound bilingual, two different lexical items represent the concept of "ordenador" and "computer", but the underlying meaning would be the same across the two languages (see Heredia & Cieślicka, 2014 for short review comparing compound-coordinate bilinguals). Lastly, *subordinate bilinguals* (Bilingual C) learn a new language with the help of a previously acquired one, so during the early stages of L2 learning, subordinate bilinguals associate L2 words with the L1 translation equivalent and access the semantics through their L1. Therefore, when a person first begins to learn a new language, the L2 is reliant on the L1 to gain understanding and meaning of unfamiliar words. For example, an English speaker learning the Spanish word "colegio" will refer to English, their L1, to translate it into "school" to gain information such as meaning surrounding that word. However, the more proficient an individual becomes, the more the two-conceptual language systems may converge into one, where one language influences the other and vice versa. No assumptions are made in terms of the conceptual system for Bilingual C, only that the L2 is reliant on the L1.

Therefore, to test the conceptual system of subordinate bilinguals Dong, Gui, and MacWhinney (2005) conducted two written experiments. The second study is of particular relevance as it was testing whether subordinate bilinguals associate L2 words with the L1 translation equivalent and access the meanings through their L1. In experiment 1, Dong et al. (2005) explored the idea of shared meanings using a classic prime-target LDT with proficient Chinese-English bilinguals, with six conditions varying in associative strength between prime and target words within languages (i.e., English - English; Chinese - Chinese), between languages (i.e., English - Chinese; Chinese - English), and in unrelated pairs as their controls. Lexical decision responses revealed significant priming effects in both the within- and the cross-language conditions. Faster priming effects were found in L1 targets compared to L2 targets, which is explained by Dong et al. (2005) to occur because the links between the L1 and conceptual representation is stronger and thus easier to access compared to the L2 and conceptual representations. If there were two separate semantic storage areas, there would not have been any cross-language priming effects present. Thus, Dong et al. (2005) concluded that the cross-language priming effects suggest that there is a shared storage system for the conceptual representations of the bilinguals' vocabularies and asymmetrical links between concepts and lexical names in the two languages.

Experiment 2 used the semantic closeness rankings (i.e., the closeness in meaning of two concepts) given by groups of Chinese – English bilinguals varying in proficiency to test for conceptual relationships that are equivalent across language translations and those that are not. For example, for the word *red*, both English and Chinese participants would rate the word to be related to the word *colour* in both languages. Whereas the word *jealousy* is more closely associated with the word *green* in English than in Chinese, since English speakers talk about "turning green with jealousy" (Dong et al., 2005; p.229) and therefore have their own conceptual system of word meanings. According to Dong et al. (2005), L2 learners tend to preserve their L1 conceptual system in the representation of L1 words, and later develop a L2 conceptual system in the representation of L2 words. The results revealed that the less proficient participants relied on their L1 to understand the L2 meaning of the word; whereas those that were proficient used meaning of words that were exclusively associated with their L2 meaning (e.g., jealousy example). Thus, Dong et al.'s (2005) findings support the idea that there is an eventual merging of languages, and that as proficiency increases, the person moves from subordinate bilingual to a compound bilingual. These findings are important because it provides evidence of the existence of different types of bilinguals and demonstrates the need for bilingual types to be considered in research as the predictions can change. For instance, if the two languages are separate, then you would expect there to be no cross-language effects to occur, however, if the languages have been integrated then the languages should affect each other. A more detailed account of the debate of storage and access in bilinguals can be found in Chapter 2.

This thesis aims to address Weinreich's (1953) types of bilinguals by ensuring that all participants are native Spanish speakers that have learnt their L2 English after the critical period. Furthermore, all bilingual participants will have similar high-proficiency levels to one another (see section 1.3 for proficiency methods used in this research programme). While controlling for proficiency cannot guarantee that the bilinguals will not associate L2 words with the L1 translation equivalent and access the meanings through their L1 Spanish, high-proficiency may push the bilingual closer to not having to use their L1. Previous research has shown that proficiency can affect language processing, and that interference of the L1 is stronger in less proficient bilinguals when processing interlingual homographs compared to those bilinguals with high proficiency (Schulpen, Dijkstra, & Schierfers, 2003; Brenders, van Hell, & Dijkstra, 2011). Therefore, any influence of the L1 Spanish will be less likely to be evident because of the bilinguals' proficiency. However, it is something that we cannot disregard as a possibility, and therefore having an influence on the results.

In addition to the different types of bilinguals, it is important to note that the language repertoire of *all* bilinguals varies over time. A distinctive feature of being bilingual is being able to make appropriate language choices based on numerous linguistic and psychosocial factors: for instance, the type of person addressed, which language to use, and how much of the other language is needed (Ritchie & Bhatia, 2013). The state of activation of the bilingual's languages and processing mechanisms has been called *the language mode* (Grosjean 1992; 2001). Grosjean (1985; 1989; 1992; 2001**)** proposed a situational continuum, where at one end bilinguals can find themselves in a bilingual language mode, in that they are communicating with bilinguals who share their languages or with whom they normally mix languages (e.g., code-switch). On the other end of the continuum, bilinguals may find themselves in monolingual language mode, in that they adopt the language of the monolingual speaker(s) or writer(s) and deactivate the other language (as much as possible; Grosjean, 2001). In addition to end points of the continuum, bilinguals can also find themselves at intermediary points, where bilinguals may be making appropriate language choices depending on context (Grosjean, 2001).

The idea of a situational continuum is mainly based around speech production (Grosjean, 1998), instead of visual processing. The main difference between auditory and visual processing is that spoken word unfolds in time, whereas written word recognition is a more organised representation and presents the visual information permanently; additionally, the reader can see where the word starts and ends. However, Grosjean (1999) did state that the situational continuum can be applied to a reading setting as bilinguals still have to make appropriate language choices to visual stimuli: "during perception, if bilingual listeners who start off in a monolingual mode determine (consciously or not) as they go along, that what they are listening to can contain elements from the other language, they will put themselves in a bilingual mode (at least partly), that is, activate both their languages (with the base language being more strongly activated)" (p.7). This is also true of readers, whether they are reading a continuous text or looking at individual lexical items interspersed with items from the other language. Simply knowing that there is a possibility that elements from the other language will be presented (in an experiment, for example) will move the bilingual away from the monolingual endpoint of the continuum (Grosjean, 2001). Just one guest word in a stream of base language words can increase this displacement. For example, in this thesis, the recruitment adverts specifically requested proficient Spanish-English bilinguals, and this may have impacted participants in that they may have acted differently in comparison to a more naturalistic setting because they might have anticipated having to use their L1 Spanish despite the task being in their L2 English. Subsequently, we may see more interference from the L1 than usual because bilinguals are anticipating having to use their L1 Spanish.

Furthermore, the language continuum is somewhat constrained by the type and characteristics of each bilingual. For example, a bilingual who understands an L2 but cannot produce written or spoken word in L2 would be considered as having a low proficiency. This type of bilingual would be unlikely to have the same monolingual mode experience as a bilingual who can both understand and produce their L2. Therefore, it is essential to keep in mind that bilinguals are diverse and to consider varying degrees of bilingualism when classifying them in research.

This thesis aims to address Grosjean's (2001) situational continuum by taking steps to ensure that participants are as close to the monolingual L2 mode as possible. Thus, all bilingual participants taking part in the experiments will have lived in the UK, their L2 (target language) environment for a minimum of three months to ensure culture familiarity (Ritchie & Bhatia, 2013). Participants will also be unaware of the researcher's bilingualism before taking part in the research. While hiding the researcher’s bilingualism is not always effective in preventing bilinguals from entering bilingual language mode (Ritchie & Bhatia, 2013), it may push the bilingual closer along to the monolingual language mode of the continuum. Therefore, any influence of the L1 will be less likely to be evident because the bilingual is prepared to use both languages at any time. However, it is something that we cannot disregard as a possibility and thus having an influence on the results.

To sum up, the literature categorises bilinguals across various dimensions; although commonly, they are classified according to the degree of proficiency and by the age of acquisition of the languages. To add to the complexity of defining how these dimensions are measured, these factors are not simply categorical constructs, they are continuous, and thus more challenging to measure. Subsequently, in the following section, there will be a discussion on how bilingualism has been measured in the literature and the implications that these differences have on cross study comparisons.

1.3 Measuring Bilingualism

While there may be no agreed definition of bilingualism, we can consider features such as language acquisition history to measure bilingual experience. However, this information pertaining to participants is sometimes insufficient; for instance, an often-cited seminal study by Caramazza and Brones (1980) on the bilingual lexicon provides no information about the participant's language background or usage for both languages. The readers are told that the Spanish – English bilinguals were native speakers of Spanish who ranged in their self-ratings of bilingual fluency from good to excellent (mean rating of 5.5 on a seven-point Likert scale). Insufficient information such as only using self-rating scales or brief language background questionnaires have also been noted in later studies (e.g., Dijkstra et al., 2010; Lemhöfer et al., 2008).

Furthermore, there is inconsistency in the literature on what bilingual characteristics are reported (e.g., Birdsong, 2016; Birdsong, & Vanhove, 2016; Hulstijn, 2012). A review of 186 studies published between 2005-2015 examined bilinguals and monolinguals in order to comprehend how bilingualism has been defined in the literature (Surrain & Luk, 2017). The authors examined how different facets of the bilingual experience are reported and revealed that usage and proficiency were the most commonly reported (79% and 77% of studies, respectively), followed by language history (67%), and the language used at school (60%). In contrast, less than half of the studies measured proficiency objectively and even fewer (30%) reported the sociolinguistic context from which the sample was drawn. The labels used to describe the bilingual groups varied and the most commonly reported label (methods used to measure bilingualism) specified the languages spoken by the participants but did not specify dominance or learning acquisition history (46%). This was followed by simply labelling the participants as “bilingual” with no additional qualifiers (31%). The least commonly used method was acquiring history of acquisition only (6%) and language dominance (2%), though, 10% of studies applied specific labels that provided some combination of the features described above. Thus, it appears that our understanding of the effects of bilingualism is limited by methodological challenges, such as studies defining and measuring bilinguals and monolinguals differently, making cross-study comparisons challenging (Surrain & Luk, 2017). Subsequently, it is crucial to clearly report and measure the characteristics that qualify participants as being bilingual (and monolingual) in research studies in order to facilitate cross-study comparison.

In this thesis, importance is placed on acquiring information about participants' language background. All participants will be adult Spanish native speakers living in the United Kingdom (UK) and be actively using both their L1 Spanish, and their L2 English. This thesis labels the bilingual group to be late bilinguals, although most participants were first exposed to the English language in primary school which was in their L1 environment. The bilinguals' performance will be compared to that of monolingual native English speakers (NEs). NEs will not be actively learning another language, would be unable to communicate comfortably in another language other than English, and have not been exposed to the Spanish language. A subjective measure of proficiency and usage of language will be used for all participants through self-rating scales for the following components: reading, writing, speaking, and understanding. This measure will form part of the background questionnaire given to all participants. In addition to the subjective measure, there will be a standardised objective measure of obtaining proficiency levels for both languages using the Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012) and the equivalent test for the Spanish language (LexTALE-ESP; Izura, Cuetos, & Brysbaert, 2014). Therefore, bilingualism was determined by the level of proficiency of bilingual speakers and measured in an objective and subjective manner ensuring consistency across the studies presented in the thesis.

The LexTALE tests use word frequency as the basic criterion for words of various degrees of difficulty; there are words that participants with low proficiency levels would know, and other words that would only be known to participants with high proficiency levels (Lemhöfer & Broersma, 2012). An example of a high frequency word is "shin", and an example of a lower frequency word and thus requiring a higher proficiency level is "wrought". The LexTALE is a test of vocabulary knowledge at a high proficiency level and includes 60 items (40 words, 20 non-words), and 90 items (60 words, 30 non-words) for the LexTALE-ESP, for which the participant must indicate whether they think a word is real in that particular language. Low scores indicate lower proficiency levels in that language, hence the higher the score the better the proficiency (Lemhöfer & Broersma, 2012). The LexTALE score consists of the percentage of correct responses, corrected for the unequal proportion of words and non-words in the test by averaging the percentages correct for these two item types. According to Lemhöfer and Broersma (2012) scores below 60% are classed as being low in proficiency, thus only bilinguals and NEs who score above this percentage were included in the research programme.

There is evidence in support of the usefulness of LexTALE in measuring language proficiency levels (e.g., Diependaele, Lemhöfer, & Brysbaert, 2013; Yap, Balota, Tse, & Besner, 2008). Diependaele et al. (2013) conducted a visual word recognition experiment and noted that participants with high scores on the LexTALE had a smaller word frequency effect than those with low scores on the test. The word frequency effect refers to the fact that reaction times (RTs) are faster for words seen more frequently than for words seen less frequently (e.g., Coltheart, 1981). These findings demonstrate the importance of controlling for proficiency levels across participants because different levels can influence results. Subsequently, the LexTALE will be used as well as self-rating scores to ensure high proficiency in both languages and consistency amongst the collected sample.

In summary, many variables, such as the way bilinguals learn an L2, the levels of proficiency of bilinguals, and the age at which the language is learned all play a role in how we define and measure bilingualism. Indeed, many factors such as the environment in which we learn a language play a role in determining whether two languages are connected in the bilingual mind (Green, 1998; Weinreich, 1953). This section has outlined that late bilinguals will be used throughout the research programme, and how and why the thesis will measure bilingualism in the form of objective and subjective measures. The next chapter will present an overview of the bilingual models relating to the debate of how bilinguals store and access their languages and, in turn, of how effective these models are in explaining how interlingual homographs are stored and accessed.

Chapter 2 – Models of the Bilingual Lexicon.

Chapter 1 introduced the construct of bilingualism, how the literature commonly measures it, and in turn how this research programme will measure bilingualism subjectively and objectively. The aim of the present chapter is to outline the main theoretical accounts that attempt to answer the two questions of storage and access in bilinguals; and explore how effectively they explain how interlingual homographs are represented in the bilingual lexicon. As mentioned previously in Chapter 1, the question around storage refers to the fact that it is debated whether all words are *stored* in an integrated lexicon, regardless of which language the word belongs to, or whether there are separate lexicons for each known language. Whereas, the question around access refers to how a bilingual can *select* the correct language to use depending on context, and whether all the languages the bilingual knows become active during lexical retrieval not just the language that is relevant. These questions are important to this research programme as it focuses on the use of interlingual homographs that introduce ambiguity between languages. Both questions require assumptions to be made about bilingual structures in the cognitive architecture. Furthermore, more evidence is needed to understand how orthographic similarity affects these questions. This will be explored in this thesis by using interlingual homographs (with identical spelling) and near-identical interlingual homographs (which differ by a minimum of one letter). The following six theoretical accounts will be discussed in this chapter in turn: the Hierarchical Model (Potter et al., 1984), the Revised Hierarchical Model (RHM; Kroll & Stewart, 1984), the Bilingual Interactive Activation Model (BIA; van Heuven et al., 1998), the Bilingual Interactive Activation Plus (BIA+) model (Dijkstra & van Heuven, 2002), the Inhibitory Control model (Green, 1998), and the Multilink Model (Dijkstra et al., 2019).

2.1 The Hierarchical Model (Potter et al., 1984) and the Revised Hierarchical Model (RHM; Kroll & Stewart, 1994).

Weinreich’s (1953) bilingual types (coordinate, compound, and subordinate) have visibly influenced much of the thinking behind the later psycholinguistic modelling of the bilingual lexicon (see Chapter 1 page 13 for a summary of the bilingual types). For instance, one of the earlier theories proposed by Potter, So, Von Eckardt, and Feldman (1984) addresses the question of storage, and how a bilingual's lexical knowledge is represented in their mind. Similarly, to what was suggested by Weinreich (1953), Potter et al. (1984) assumed that some type of connection between the native language (L1) and second language (L2) must be made during the learning of new L2 words. Based on this assumption, Potter et al. (1984) proposed The Hierarchical Model and put forth two main hypotheses, the concept mediation and word association hypotheses, as possible theories that could explain how connections between L1 and L2 are made. The concept mediation hypothesis suggests that each language has independent access to a common conceptual system. This hypothesis follows the structure of Weinreich's (1953) coordinate bilingual (see Figure 2). In addition, the word association hypothesis suggests that the L1 and L2 are connected at the lexical level, but that only the L1 has access to the conceptual system. That is, words in the L2 are understood through L1 lexical representations. This hypothesis follows the structure of a subordinate bilingual. Potter et al. (1984) also proposed a third hypothesis that mirrors the theoretical transition from subordinate to coordinate bilingual. The intermediate hypothesis suggests that while learning L2 vocabulary, L2 learners first acquire lexical associations through their L1, hence performing under the word association hypothesis. Gradually, direct links are developed between the L2 words and their conceptual representations as in the concept mediation hypothesis.

To test these three alternative hypotheses, Potter et al. (1984) tested proficient Chinese-English and non-fluent French-English bilinguals across three tasks: picture naming in both languages, word reading in the L1, and lastly, translating words from L1 into the L2. Predictions of how participants would respond to words and pictures when using L2 varied depending on which hypothesis is selected. For instance, the word association hypothesis predicts that translating from L1 into L2 would involve 3 processing steps: recognition of the L1 word, retrieval of the L2 word, and speaking the L2 word. See Figure 3 on page 27 for visual depiction of the model. For the picture naming task in a L2, there would be an additional two processing steps and the sequence would be as follows: picture recognition, concept retrieval, retrieval of the L1 word, retrieval of the L2 word, and speaking the L2 word (see Figure 3). Word and picture recognition are assumed to be approximately equal (i.e., T/S in Figure 3), and this is supported by the recognition data for both stimuli in Potter et al.'s (1984) study. However, due to the extra 2 steps in the picture naming task, it is predicted that it would take more time to complete it than the word translating task (i.e., T + V + U + W + Z versus S + W + Z).

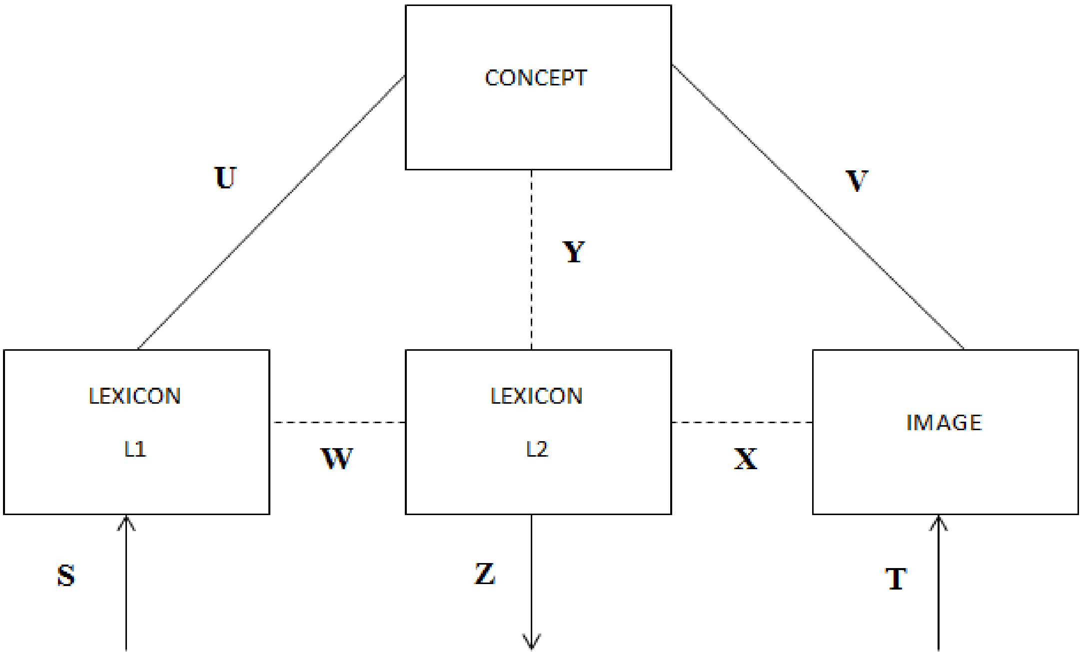

Figure 3.
An illustration of the Hierarchical Model's (Potter et al., 1984) major components for picture and word processing. Source: Adapted from Chen and Leung (1989).

In figure 3 the letters represent processing times for different mental steps: **S** = recognition of a L1 word, **T** = recognition of a picture, **U** = concept retrieval from the L1 lexicon, **V** = concept retrieval from the image, **W** = retrieval of a L2 word from the L1 lexicon, **X** = retrieval of a L2 word from the image, **Y** = retrieval of a L2 word from the concept, and lastly, **H** = speaking a L2 word.

In contrast to the word association hypothesis, the concept mediation hypothesis predicts that the translation and the picture naming task would involve similar processing steps. That is, for the translation task, the individual would first recognise the L1 word, followed by concept retrieval, retrieval of the L2 word, and lastly, speaking the L2 word. Similarly, for the picture naming task, the individual would recognise the picture rather than the word, followed by concept retrieval, retrieval of the L2 word, and lastly, speaking the L2 word. Reaction times for semantic access for pictures and words are assumed to be equal (i.e., T + V / S + U), this assumption is supported by the categorisation data from Potter et al's (1984) study. Therefore, translating and picture naming in a bilingual's L2 should take approximately the same time to complete (i.e., S + U + Y + Z versus T + V + Y + Z). The results demonstrated that no significant differences were found in either of the bilingual groups when responding in their L2 between picturing naming and word translating. Subsequently, the authors concluded that L2 words are associated with corresponding words in L1 via a common conceptual system, even for non-fluent bilinguals, not by means of direct associations between vocabulary items.

Although the data from Potter et al.'s (1984) research supported the concept mediation hypothesis, Chen and Leung (1989) argued that this may be the case because the bilinguals used in Potter et al.'s (1984) research had different amounts of training in their L2. For example, the English - French bilinguals had 2-3 years of training in high school and already knew enough French to enable them to spend a summer in France (see Potter et al., 1984, p. 32); whereas the Chinese-English bilinguals had at least 12 years training in school. Consequently, Chen and Leung (1989) argue that it is not clear whether the age of acquisition (AoA) of the languages and amount of formal training in the L2 could have contributed to the lack of differences between the tasks. Therefore, to address this, Chen and Leung (1989) proposed three experiments, using similar tasks to Potter et al. (1984; reading words aloud, picture naming, and word translating), to investigate the possible effects of AoA of the languages and L2 proficiency on an individuals' pattern of lexical processing. There were three bilingual groups in the study: proficient Chinese-English adults (mean age of 20 years; studying their L2 for approximately 12 years in school), low proficient Chinese-English adult beginners (mean age of 20 years; studying their L2 for approximately 2 years), and low proficient Chinese-English children beginners (mean age of 7 years; studying their L2 for approximately 2 years). The results revealed that all bilingual groups made fewer errors and were quicker in reading words compared to naming pictures when responding in their L1. When the task was in their L2, the proficient bilinguals were equally efficient in both the translating word and picture naming tasks. However, the adult beginners were faster in the word translation task than the picture naming, whereas the child beginners showed faster RTs in the picture naming task compared to the word translation. These findings are consistent with the idea that proficient bilinguals can directly access the meanings of words in their L2; whereas adult beginners tend to use L1 words and child beginners use pictorial representations. Therefore, the results provide support for the word association and intermediate hypotheses and suggest that proficiency and AoA of an L2 can influence the pattern of lexical processing in the L2.

Furthermore, Chen and Leung (1989) demonstrated that there are observed asymmetries in translation performance by late bilinguals who acquired their L2 after early childhood and for whom the L1 remained the dominant language; this observation is consistent with later studies (e.g., Sholl, Sankaranarayanan, & Kroll, 1995). Therefore, to account for the asymmetries in translation, the Hierarchical Model (Potter et al., 1984) was revised a decade later by Kroll and Stewart (1994) and referred to as the Revised Hierarchical Model (RHM). The RHM accommodates the proposal that L2 processing sometimes shifts from its reliance upon L1 connections for meaning to direct conceptual connections, suggesting that an asymmetry may exist in the strength of lexical-conceptual connections between languages in the bilingual (Kroll & De Groot, 1997; 2002). The RHM essentially merged the concept mediation and word association hypotheses into a single developmental model.

The RHM model (Kroll & Stewart, 1994), like the original, does not support the idea of an integrated lexicon for bilinguals but accepts direct associative links between translation equivalents. Furthermore, this model suggests that when a bilingual learns an L2, lexical connections are formed between the two languages. As a bilingual becomes more fluent in the L2, stronger direct links are established between the L2 and the conceptual system. However, these links would not cause the already existing lexical links from the L2 to the L1 to disappear. Thus, according to the RHM (Kroll & Stewart, 1994) a distinction had to be made between two types of word representations: lexical and conceptual representations. At the *lexical representation level* (which contains information about word forms) two lexicons are distinguished, one for words of the L1 and one for words of the L2. The lexical store for the L1 is bigger compared to the L2 store because even for highly proficient bilinguals, it is assumed that more words are known in the L1 than in the L2. Translation equivalents in the two lexicons are linked via excitatory links and are connected to a shared *conceptual representation system* that contains the meaning of the words (see Figure 4).

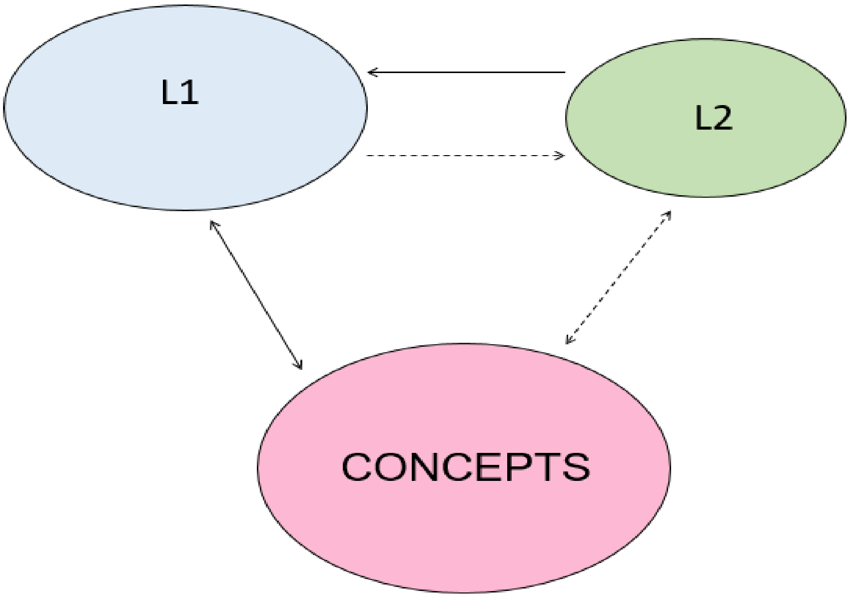

Figure 4.
The Revised Hierarchical model (Kroll & Stewart, 1994) in pictorial form. The arrows show the lexical links between the languages and concepts. Source: Adapted from Kroll and Stewart (1994).

Potter et al. (1984) put forward the explanation that longer translation RTs from L1 to L2 (forward translation) compared to those from L2 to L1 (backward translation) are due to an underlying asymmetry in the strength of the links between words and concepts in each of the bilingual's languages. Connections between L1 words and concepts are assumed to be stronger than those between L2 and concepts because the L1 has an advantaged access to meaning. Whereas the L2 is more likely to require mediation via the L1 translation equivalent until the bilingual acquires sufficient skills in the L2 to access meaning directly. On this account, translation from L2 to L1 can be accomplished lexically, without semantic access, if the L2 word enables lexically mediated retrieval of the translation. In contrast, L1 to L2 translation would be more likely to engage semantics because of the strong L1 link to meaning (e.g., Kroll & Stewart, 1994; Sholl et al., 1995).

The RHM (Kroll & Stewart, 1994) has played a key role in how bilingual memory thinking has developed. One of the strengths of the RHM is that it seemed to support earlier views that considered bilingualism as the co-existence of two independent language systems. For example, one of the first studies to investigate interlingual homographs was conducted by Gerard and Scarborough (1989) who tested Spanish-English bilinguals in a two-part LDT that used both identical cognates and interlingual homographs. Participants made lexical decisions to targets that were primed to be unrelated to either language of the cognate and interlingual homograph or related to one of the meanings of the interlingual homographs in either language. In part 1 of the experiment, half of the bilinguals made lexical decisions in their L1 Spanish (Spanish targets), and the other half of bilinguals in their L2 English (English targets); and in part 2 this was reversed, so participants who first made lexical decisions in their L1 Spanish would now be making lexical decisions in their L2 English. In the first part of the experiment, bilinguals in the English target condition showed no influence of their knowledge of Spanish when making lexical decisions to cognates or interlingual homographs because there were no differences compared to matched control words. This was also true for Spanish target bilinguals who showed no influence of their knowledge of English when making lexical decisions. This finding would therefore fit with the RHM’s account of the bilingual lexicon being organised into separate lexicons. The independent organisation of languages provides a straightforward explanation of why these bilinguals do not show interference from the non-target language, while processing the target language.

In part 2, however, participants reacted faster to cognates and interlingual homographs compared to control words which indicated facilitative priming. Gerard and Scarborough (1989) argue that this is because the identical spelling in cognates and interlingual homographs allows for the facilitation of encoding, thus making it easier to process these words for bilinguals at the conceptual shared level (see concepts in Figure 4 on page 30). Whilst the findings from part 2 do not provide strong evidence for selective access, Gerard and Scarborough's (1989) analyses of the frequency of words does. Specifically, words with higher frequencies in the target language were responded to faster than lower frequency words. For example, in the case of interlingual homographs, bilinguals were slower to recognise low-frequency words in English such as fin (meaning *end* in Spanish), in comparison to high-frequency words in English such as red (meaning *network* in Spanish). When the same words appeared in the Spanish-target condition, the opposite pattern was found; now bilinguals were quicker to recognise the word fin which has a high-frequency in Spanish, and slower to recognise words such as red with a lower-frequency in Spanish. The authors attribute these frequency effects to the suggestion that lexical information is stored in functionally separate language-specific lexicons; and according to the selective access views, bilinguals can inhibit or activate one of their languages, depending on context (see van Heuven et al., 1998 for a discussion of the different possible organisations of a bilingual visual word recognition system). Therefore, Gerard and Scarborough (1989)'s results suggest that even as task demands change (i.e., target language) bilinguals can exhibit some level of control over their languages (see Inhibitory Control Model subsection 2.3 for more information about bilingual control mechanisms; Green, 1998).

However, in contrast to Gerard and Scarborough's (1989) research which supports the structure of the RHM (Kroll & Stewart, 1994), the model has been challenged by subsequent findings in bilingual language processing research (for a review, see Brysbaert & Duyck, 2010). In their review of the RHM, Brysbaert and Duijck (2010) presented further evidence of “L1 and L2 words acting very much as if they are words of the same language, interacting with each other as part of the word identification process” (p. 364). In particular, strong evidence in favour of an integrated lexicon is data which demonstrates that orthographic neighbours (ONs; an ON is a word of the same length that differs from the original string by only one letter) of the other non-target language have an effect on target word recognition in the target language (van Heuven et al., 1998). van Heuven et al. (1998) used proficient Dutch-English bilinguals to investigate how the recognition of target words belonging to the L2 (English) was affected by the existence of ONs from each language. Increasing the number of ONs in the L1 slowed RTs and produced inhibitory effects, while an increase in ONs in the L2 produced facilitatory effects for target words. Monolingual NEs also showed facilitation due to an increase in ONs, but no effect in Dutch ONs which is logical as there is no Dutch knowledge to draw upon. The results suggest parallel activation of words in an integrated lexicon because of cross-linguistic word form effects that are difficult to explain using the separate lexicon hypothesis supported by the RHM (Kroll & Stewart, 1994). One could suggest that near-identical interlingual homographs act as an interlingual ON, as they are lexically related to the target word (e.g., if the near-identical interlingual homograph target word is CARPET, then the lexically related word is CARPETA which is the near-identical equivalent and means *folder* in Spanish).

Indeed, in the years since the review (Brysbaert & Duijck, 2010), further evidence has been collected showing, for instance, that word candidates which are morphologically related to a target word are activated even when they belong to another language. For example, when English-Dutch bilinguals read the English word *house*, they may activate an English word such as *housekeeper* and also a Dutch word such as *werkhuis* (work house; Mulder, Schreuder, & Dijkstra, 2013; Mulder, Dijkstra, Schreuder, & Baayen, 2014). Despite these findings suggesting an integrated lexicon, the RHM (Kroll & Stewart, 1994) is still firm in its stance of bilinguals having independent lexicons as Kroll, Van Hell, Tokowicz, and Green (2010) argue that "it could very well be the case that the two lexicons are functionally separate but with parallel access and sub-lexical activation that creates resonance among shared lexical features" (p. 374). Thus, they suggest that these shared features are the reason why parallel activation is found, not because of an integrated lexicon.

This notion of shared features at the conceptual level could also explain why later studies using interlingual homographs as their stimuli in lexical decision and priming tasks have found inhibitory or no effects at all (e.g., De Bruijn, Dijkstra, Chwilla, & Schriefers, 2001; De Groot, Delmaar, & Lupker, 2000; Dijkstra, De Bruijn, Schriefers, & Ten Brinke, 2000a; Dijkstra, Grainger & van Heuven, 1999; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Dijkstra, Timmermans, & Schriefers, 2000b; Lemhöfer & Dijkstra, 2004; Libben & Titone, 2009; Macizo, Bajo, & Martín, 2010; Poort et al., 2016; Titone, Libben, Mercier, Whitford, & Pivneva, 2011). The RHM (Kroll & Stewart, 1994) can explain why no effects are found when processing interlingual homographs because it assumes separate lexicons for each language, and hence, no effects are present as they do not interfere with one another. However, it is also able to explain why semantic interference effects are present through the underlying sharing of features of the interlingual homographs (Luo, Craik, Moreno, & Bialystok, 2013), though the meaning differ across languages, all meanings are stored at the conceptual level, and there are lexical links between the lexical and conceptual level which can lead to interference effects when processing interlingual homographs.

In contrast to the RHM, the next model that will be outlined argues that cross-language effects such as those seen with interlingual homographs are due to an integrated lexicon.

2.2 Bilingual Interactive Activation (BIA; van Heuven et al., 1998) and Bilingual Interactive Activation Plus (BIA+) Model (Dijkstra & van Heuven 2002)

The models discussed above provide a plausible description of how the bilingual lexical organisation and access might occur, though, they do not provide a detailed account of how lexical access occurs in bilinguals. The Bilingual Interactive Activation Plus (BIA+) model, on the other hand, does incorporate lexical access and was proposed by Dijkstra and van Heuven (2002) as a revised version of the localist-connectionist Bilingual Interactive Activation (BIA) model (van Heuven et al., 1998). In turn, the BIA model was a straightforward extension of the Monolingual Interactive Activation model of visual word recognition, first proposed by McClelland and Rumelhart (1981). The model contained the three layers used by McClelland and Rumelhart (1981) which are features, letters and words, and to that the BIA model added L2 words to the word layer and a language layer at the top.

The BIA and BIA+ models involve orthographic, phonological, and semantic representations that are connected both within and between languages (van Heuven et al., 1998; Dijkstra & van Heuven, 2002). In contrast to the RHM, the BIA and BIA+ models both assume that lexical representations across languages are stored in an integrated lexicon. Moreover, the BIA and BIA+ models propose that the visual presentation of a word leads to co-activation of many word candidates from different languages that are similar to the target input. For instance, when English - Dutch bilinguals see a word, the word activates both the orthographic and phonological neighbourhoods of both languages. The BIA model is related to orthographic information only, whilst the BIA+ includes both phonological and orthographic information. Therefore, in the case of the BIA+ model for the example target word *HOOD*, orthographic activation will include English neighbours such as *food* and *hold*, and Dutch neighbours such as *hond* (translated to dog in English) and *lood* (translated to lead in English). These orthographic representations will then begin to activate their meaning representations (e.g., Balota, 1994; Grainger, 2008). Subsequently, the meaning of *food* and *hond* will be theoretically activated to some extent in parallel with the meaning of the target word *HOOD*. Furthermore in the BIA+ model (Dijkstra & van Heuven, 2002), active orthographic representations also activate their phonological neighbours (PNs; e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). According to Luce and Pisoni (1998) PNs refer to the number of words that vary in their phonological structure from another word based on a single phoneme that can be substituted, deleted, or added. In addition to orthographic and phonological neighbourhoods, semantics also play an important role in the activation of representations and disperse their activation. For instance, in the example of the target word *HOOD*, this may spread activation to the semantic meanings of *car*, and to the ON *food* which may lead to semantic activation of *hungry*. Possible word candidates are activated; however, the task system continuously churns out related/possible candidates from the word identification system to eventually make a response relevant to the task.

To conceptualise this process the BIA and BIA+ identified two interactive subsystems: word identification and task/decision (task schema). According to the updated BIA+ model (see Figure 5 on page 38), during word identification, the visual input activates the sub-lexical orthographic representations simultaneously activating the orthographic whole-word lexical and the sub-lexical phonological representations. That is, as a letter string is processed, orthographic and phonological features of a word activate words of similar orthography and phonology in parallel, which in turn interact with semantics (as seen in the *HOOD* example before*)*. A final language identification node, not previously defined in the *HOOD* example, contributes to the activation process by interpreting language-specific cues which indicate language membership (i.e., to which a word belongs; Dijkstra & van Heuven, 2002). All this information is then used in the task/decision system (task schema), which decides which actions must be performed based on the relevant information that becomes available after word identification processing (van Heuven & Dijkstra, 2010). The task schema system was inspired by Green's (1998) inhibitory control model that will be discussed in the following sub-section.

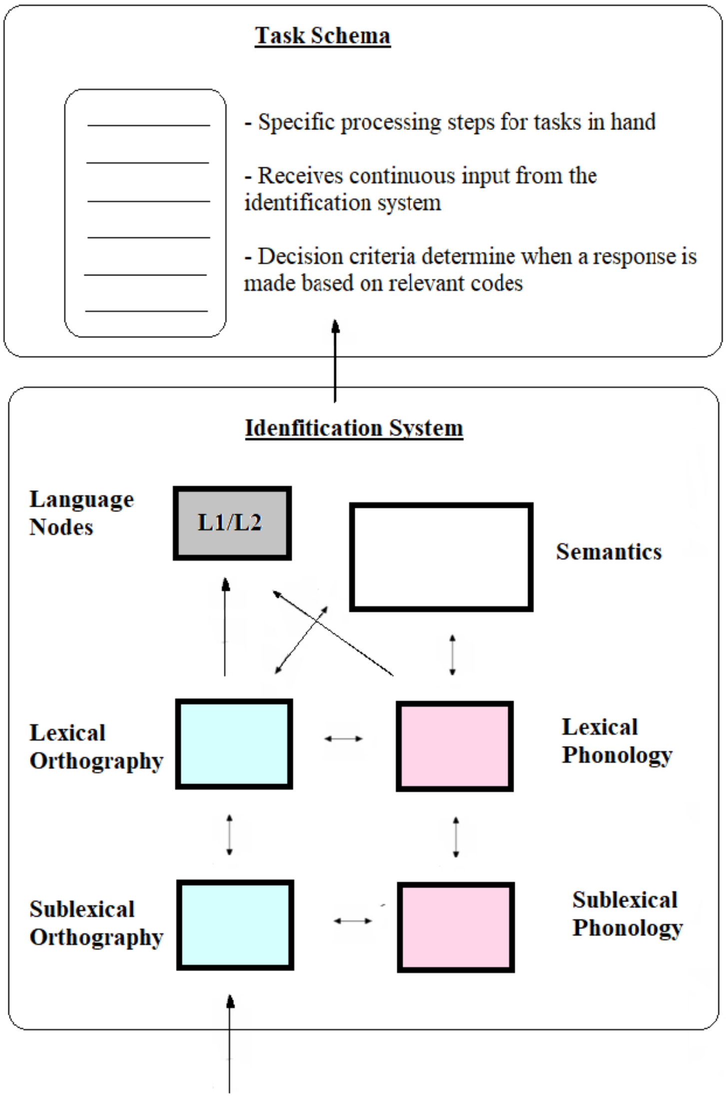

Figure 5.

BIA+ Model (Dijkstra & van Heuven, 2002) in pictorial form. Arrows show activation and movement between representational pools. Non-linguistic context only affects the task schema level. Source: Adapted from Dijkstra & van Heuven, 2002.

While the BIA and the BIA+ share similarities, there are a few distinct differences between the two. For example, in the BIA model, language nodes play an inhibitory role, whereas, in the BIA+ model they are assumed to not affect the activation levels of the word identification system. The BIA+ model positions itself under the assumption that language nodes exist to provide a representation for the language of membership, and this is based on the information from orthographic and phonological processes. These language nodes, however, have been found to have no effect on the activation level representation of words, because the focus of activation of these nodes is post-lexical (i.e., the word has already been activated before reaching the language node; van Heuven & Dijkstra, 2010). Subsequently, the existence of these nodes allows bilinguals to keep interference from the non-target language to a minimum, while still allowing bilinguals to process the target language. Nevertheless, interference of the non-target language does still occur as evidenced by studies with interlingual homographs (e.g., Macizo et al., 2010; Poort et al., 2016; see chapter 4 for more detail about these effects), and thus the presence of a language input produces activation of items (e.g., ONs; PNs) that overlap with the target input in either language, suggesting an integrated lexicon.

Within the BIA+'s theoretical framework, the processing of interlingual homographs can be understood by assuming that they are represented by two, with the possibility of partial overlap, competing representations which interfere with each other (Dijkstra & van Heuven, 2002; van Heuven & Dijkstra, 2010). This is because both readings of the interlingual homograph have separate orthographic (with the possibility of partial overlap) and semantic nodes (Dijkstra et al., 2010; Peeters et al., 2013). Moreover, in the BIA+ all words have a resting activation level that relies on the frequency usage and the similarity of the target word to the input, so that all possible candidates are activated until one is selected. That is, after a period of competition, a word is recognised when it passes through the recognition threshold. The orthographic nodes of words inhibit each other through the process called lateral inhibition, irrespective of the language to which they belong and environmental language context. The task system continuously churns out different activations (e.g., related/possible candidates) from the word identification system which weighs up the different levels of activation to eventually make a response relevant to the task. The more frequent a candidate or the more similar to the target, the faster the processing.

The inhibition effects seen with interlingual homographs (see chapter 4 for discussion of these effects) are explained by the BIA+ model (Dijkstra & van Heuven, 2002) as a disproportionately strong effect of lateral inhibition. As with any other two words, the orthographic nodes of the interlingual homograph laterally inhibit each other. Although, because these representations are identical (or near), competition is stronger compared to non-ambiguous words. Therefore, it is proposed that both meanings of an interlingual homograph are activated non-selectively until a relevant meaning is selected. Whilst some control processes are needed for adequate lexical selection and task-related actions, non-selective access is consistent with research showing that the non-target language in interlingual homographs (and cognates) interferes and is activated in tasks that are exclusively undertaken in one language (see Lauro & Schwartz, 2017 for review). Moreover, both models only simulate the orthographic recognition of 4-5 letter words. Consequently, this limits the generalisability of research studies where the stimuli are above this letter length. These weaknesses are particularly relevant for the current research programme, as some of the stimuli used are above the 5-letter length (see Appendix 1 for full list of interlingual homographs used in this thesis) and therefore, would make it hard to explicitly account for the processing of all interlingual homographs used in this research.

2.3 Inhibitory Control (IC) Model (Green, 1998)

Another model which is well cited in the literature is the IC model proposed by Green (1998), that, though it makes limited claims about the organisation of the bilingual lexicon, it focuses on the importance of task demands and regulation that happen during language processing by modifying the levels of activation of items in a language network (Green, 2003). As we have seen, in the BIA+ language nodes are not successful in completely inhibiting a particular language and are therefore not able to prevent interference from the non-target language. In the IC model however, language schemas (similar to language nodes) are separate from the lexical semantic system (see Figure 6). The language task schemas are moderated by a supervisory attentional system that regulates their activity. By inhibiting the task schema for the non-target language, it can limit interference and maintain activation for the target language. However, to switch languages and reactivate an inhibited language, latent inhibition must be overcome, and this is linked with a processing cost (Meuter & Allport, 1999). The IC model is based around speech production but could be applied to a written word comprehension context, as inhibition of the non-target language is still required when reading words.

A core assumption of the IC model is that language production is a communicative action that is equivalent to non-linguistic physical actions (Green, 1998; Abutalebi & Green, 2007; see Figure 6). Like physical actions, language production comprises of mental task schemas, which are action sequences that are employed by a conceptualizer. These task schemas achieve particular goals, and for any given goal, parallel activation of multiple task schemas compete to control output. Subsequently, the supervisory attentional system (SAS; Schallice & Burgess, 1996) suppresses goals that are not necessary via inhibitory control operations and monitors the successful implementation of current goals, based on input from the bilingual lexico-semantic system. For example, when a bilingual engages with a monolingual speaker in their L2, the conceptualizer relays input from the bilingual lexico-semantic system to the SAS, which, in turn, implements greater inhibitory control to suppress the irrelevant but more routinely used L1 dialogue language schema (see Figure 6).

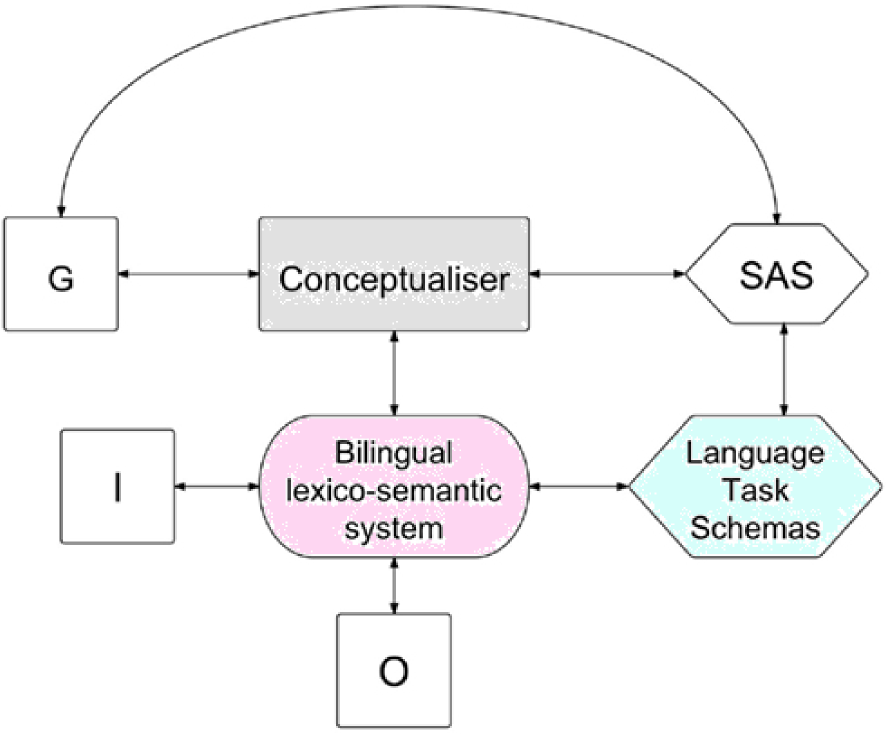

Figure 6.
Inhibitory Control Model (Green, 1998) of bilingual language production. G (goals); I (input); O (output); SAS (supervisory attentional system). Source: Adapted from Green, 1998.

Multiple studies have tested this model in the domain of language production (for a review of studies related to language production, see Kroll & Gollan, 2014), but by far fewer in the domain of language comprehension (e.g., Macizo, Bajo, & Martín, 2010; Martín, Macizo, & Bajo, 2010; Mercier, Pivneva, & Titone, 2014; Pivneva, Mercier, & Titone, 2014). According to the IC model, the amount of inhibition used to resolve interference in language comprehension should match the level of activation of a given language (Abutalebi & Green, 2007; Abutalebi, & Green, 2008). In other words, the more proficient a person is in a language, the more inhibition is required to supress the activation of it. For instance, when a bilingual uses their L2, candidates from the more dominant L1 are strongly activated and compete with the L2 target candidates. Furthermore, the more strongly a representation has been inhibited, the longer it takes to overcome inhibition and recover its normal activation level when the representation needs to be accessed again. Whilst previous research on the relationship between L2 proficiency and the cognitive mechanisms underlying the resolution of co-activation between languages is highly interesting (for more information on this relationship, see Durlik, Szewczyk, Muszyński, & Wodniecka, 2016), it unfortunately falls outside the scope of the current research programme as no manipulation of proficiency is included and only proficient bilinguals are used, although the possible involvement of inhibition in resolving interference will be tested in the experimental studies.

However, Macizo et al. (2010) did investigate whether using the less dominant L2 required more inhibition of L1 when the targets were related to the semantic meaning or were a direct translation of the interlingual homograph. Spanish – English bilinguals carried out a written semantic relatedness judgement task where they had to decide whether pairs of L2 words were or were not semantically related to each other. In order to maintain a L2 context and not activate the L1 Spanish, all stimuli and instructions were presented in English and participants were not told about the presence of interlingual homographs in the task. The task consisted of two-pair blocks. In Pair 1, the experimental condition was an interlingual homograph presented alongside a second L2 word which was related to the Spanish L1 meaning of the interlingual homograph and unrelated to the L2 (e.g., PIE-TOE). Pair 1 control pairs, instead, consisted of two unrelated English words (e.g., LOG-TOE). Hence, the participant's correct response to Pair 1 trials should be *no*, although in the bilinguals the incorrect response *yes* was misleadingly primed by the L1 Spanish reading of the interlingual homograph (*foot*). The results demonstrated that participants needed significantly more time to respond to pairs that included interlingual homographs and the Spanish related meaning compared to control pairs, and they produced over three times as many errors than in pairs consisting of only L2 English words. Macizo et al. (2010) argued that these findings were evidence of co-activation and interference between languages, and therefore, evidence of language non-selective access. In Pair 2 trials, the experimental condition consisted of an L2 English translation of the L1 Spanish interlingual homograph meaning (e.g., *foot* as a translation of the Spanish meaning of PIE) presented alongside a related English word (e.g., HAND). Control pairs consisted of two related English words in which the translation was replaced with another English word related to the second word in Pair 2 (e.g., FINGER). Macizo et al. (2010) were particularly interested in the RTs and accuracy in the translation versus the control condition in Pair 2 when that was preceded by the interlingual homograph condition in Pair 1. The results revealed that in the trials following the interlingual homograph condition in Pair 1, participants had slower RTs and were less accurate to respond to pairs containing interlingual homographs translations than to the control pairs. No differences were observed between conditions in Pair 2 for those trials that followed the control condition in Pair 1. Macizo et al. (2010) argued that the observed interaction resulted from inhibition of the interlingual homographs' L1 irrelevant meaning in Pair 1, which had to be reactivated in Pair 2. Therefore, Pair 2 provides an index of inhibitory control over competing meanings of the interlingual homographs. This pattern of results has been successfully replicated by authors in follow-up studies in the Polish language (Durlik et al., 2016) and French language (Pivneva et al., 2014) giving greater validity and generalisability to Macizo et al.’s (2010) findings.

In sum, the IC model appears to effectively describe a method by which language activation can be adjusted, suggesting that the initial conflict between two languages, both in production and in comprehension, is resolved by active inhibition. However, the efficiency by which the inhibitory demands are managed may vary between different types of bilinguals; hence, not all bilinguals will be influenced by inhibitory processes to the same extent (Durlik et al., 2016). Therefore, to control this confounding variable, the current thesis will include late bilinguals as defined in chapter 1 with similar proficiency levels, to capture consistent and similar inhibition levels.

2.4 Multilink Model (Dijkstra et al., 2019)

The last model to be discussed in this chapter is the Multilink model which was proposed by Dijkstra et al. (2019) and is centred on the principle that "no model should be left behind" (Kroll, Van Hell, Tokowicz, & Green, 2010, p. 379). As such, the Multilink model can be considered a hybrid of various previous models as it implements the basic structure of the RHM, consisting of the input and output language systems connected via shared meanings; and the task/decision system of the BIA/BIA+ model. The standard network architecture of the model can be seen in Figure 7.

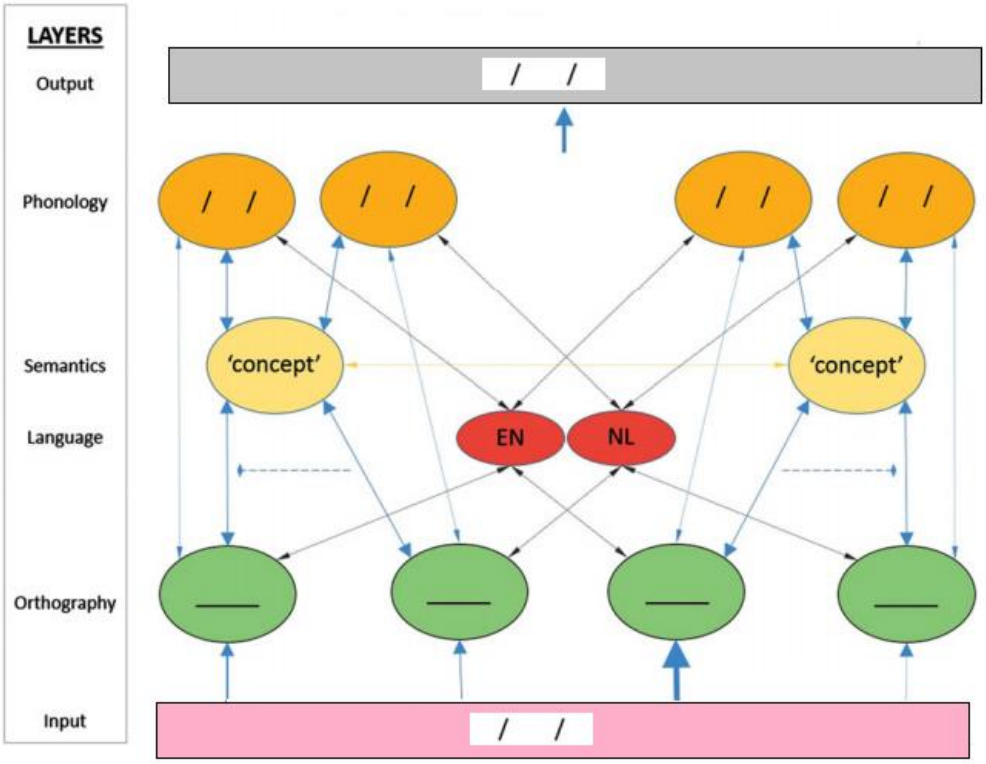

Figure 7.

The architecture of the Multilink model (Dijkstra et al., 2019). Source: Adapted from Dijkstra et al. (2019).

In Figure 7 the input is indicated in the pink box with dashes (/ /), orthographic representations in the green circles with an underscore (_), phonological representations are represented in the orange circles with dashes (/ /). EN represents English language and NL the Dutch language in the red circles. The blue dashed line between two connections from orthographic to semantics indicates that their activation is summed after taking half of the second nodes activation input. Dashes indicate phonological output in the same or a different language.

According to the Multilink model (Dijkstra et al., 2019), when a person sees a written word, it accesses various lexical-orthographic representations that are consistent with it, which in turn activate their semantic and phonological counterparts, in addition to associated representations related to language membership. The potential activation of either language occurs because Multilink model assumes that words are represented in an integrated lexicon, also referred to as the base or enriched lexicon (see Dijkstra et al., 2019 for more information of simulations), with language non-selective access and parallel activation of neighbours. Similarly, to the BIA+ model (Dijkstra & van Heuven, 2002), activation of words is dependent on various factors such as their orthographic similarity to the input word and their frequency of use which will now be defined. Levenshtein distance is used as the measure of orthographic similarity (see Schepens, Dijkstra, & Grootjen, 2012), and is described as the minimum number of insertions, deletions, or substitutions required to change one word into the other (Levenshtein, 1965; 1966). Activation of a word is influenced by its resting level activation (RLA) which is based indirectly on its frequency. Dijkstra et al. (2019) argue that the RLA depends on the rank of each word after sorting all words in the lexicon from high to low frequency. However, there are limitations to putting a ranking system in place, as it does not address the differences between two words which are equal in rank, as the step size between such words in the list remains the same across lexicons of the same size. Furthermore, because the size of the bilingual's lexicon determines the step size between the frequency ranks, when a bilingual learns a new word, it adds a word to the lexicon resulting in a change in step size, and therefore impacts all the words in the lexicon. These limitations of the RLA have been addressed by the authors and more information on their log-transformations can be found in their paper (Dijkstra et al., 2019).

However, despite the log-transformations of the frequency data, a solution is yet to be established.

The proposers of the Multilink model (Dijkstra et al., 2019) consider it to be interactive, whereby all activation flows are bidirectional. However, at present, semantic representations are simple holistic units, thus, spreading of semantic activation between associated representations is left currently unconsidered (Dijkstra et al., 2019). Moreover, whilst Figure 7 shows the lexical network through which activation flows, it does not represent the task/decision system that selects particular representations for output and specifies responses depending on the task and stimulus list. For example, for a LDT, the task/decision system may check the language membership of input and output, and the degree of orthographic, semantic, or/and phonological activation as requirements for a response to take place. Therefore, at present, the model assumes that lexical decisions for words take place based on activation thresholds of lexical orthographic representations. Furthermore, similar to the BIA+ model, the Multilink model makes the assumption that activation in the word retrieval system flows in the same way even in different task situations. Differences in patterns of findings are considered to originate and be influenced by task demands, parameter settings, and use of both linguistic and non-linguistic context. Dijkstra et al. (2019) argue that the model can be improved by including a task/decision component, such as those that exist between LDT and word naming tasks (De Groot, Borgwaldt, Bos, & Van den Eijnden, 2002), to take into account task differences and make the model more adaptable.

In their review, Costa and Pickering (2019) generally conclude that the Multilink is a useful model. However, they focus on the absence of an explanation of how L2 learning develops, and the consequences that this may have for the architecture of the L1 and L2 lexicon. Though, this notion can be applied to all models, as it is important to start capturing how learning unfolds and the consequences of such learning. In Chapter 1, we were introduced to how different types of bilinguals could emerge as a result of how the languages are learnt, and yet, it appears this notion of learning has not continued into later models. However, Costa and Pickering (2019) do note that the Multilink authors are aware of this gap in the model as they briefly discuss it (Dijkstra et al., 2019).

On the other hand, compared to the BIA (van Heuven et al., 1998) and BIA+ (Dijkstra & van Heuven, 2002) models' letter length limitations, the multilink model does simulate the recognition of 3-8 letter words. In fact, in Multilink simulations, input target words of *any* length activate stored orthographic word representations, potentially from several languages. Nonetheless, whilst the Multilink model addresses the processing of cognates, it currently does not provide explanations for the processing of interlingual homographs, which are the key stimuli of the current thesis. For cognates, the Multilink model explains that, because cognates consist of two identical orthographic nodes and a shared semantic node, the facilitation effect is a result of the semantic node receiving extra activation from the two identical orthographic nodes (Dijkstra et al., 2019). However, for interlingual homographs the current model does not include a lateral inhibition mechanism as the parameter has been set to 0, so words do not compete with each other and word recognition simply occurs when a word reaches the recognition threshold. Thus, unlike the BIA+ model (Dijkstra & van Heuven, 2002), which explains any interlingual homograph effects through lateral inhibition, the Multilink model is yet to offer an explanation for how interlingual homographs are processed and for how interference from competing meanings is resolved (Dijkstra et al., 2019). For this reason, the Multilink model is of limited use in explaining the results of the studies presented in later chapters and therefore findings will be interpreted according to the assumptions of the BIA+ model (Dijkstra & van Heuve, 2002).

2.5 Chapter summary

From the models discussed above, it appears that the BIA+ model (Dijkstra & van Heuven, 2002) offers the only detailed account of the representation of interlingual homographs in the bilingual mental lexicon. The two semantic readings of an identical interlingual homograph share none of their representations as both readings have separate orthographic and semantic nodes (Kerkhofs et al., 2006), and this is also true for near-identical interlingual homographs. The BIA+ model (Dijkstra & van Heuven, 2002) explains interlingual homograph effects as a disproportionately strong effect of lateral inhibition. As with any two words, the orthographic nodes of the interlingual homograph laterally inhibit each other. However, because these representations are identical, this competition is stronger than that between two regular words. In contrast to the BIA+ model, the Multilink model (Dijkstra et al., 2019) currently has its lateral inhibition set to 0, and therefore struggles to explain how and why competition is resolved in the case of an interlingual homograph. None of the models above comment on near-identical interlingual homographs, though one could make predictions similar to those for their identical counterparts and explain any effects through the process of lateral inhibition.

This thesis will provide insight to whether L1 activation is present during identical and near-identical interlingual homograph processing in an isolated and sentence context paradigm in the bilinguals’ L2. If the non-target L1 Spanish is activated it more closely supports the theoretical perspective of an integrated lexicon which is in line with the BIA (van Heuven et al., 1998) and BIA+ model (Dijkstra & van Heuven, 2002); and while the Hierarchical Model (Potter et al., 1984) and the RHM (Kroll & Stewart, 1994) opt for a separate lexicon, L1 interference effects can still be explained at the conceptual level. However, if L1 characteristics do not appear to be activated during the L2 task, then it would favour the notion of separate lexicons as suggested by the RHM, and according to the IC model (Green, 1998) that the bilingual enforced control over their non-target language.

The next chapter will introduce a novel set of identical and near-identical Spanish - English interlingual homographs with their linguistic characteristics and will investigate whether identical and near-identical interlingual homographs are processed in a similar or different way from one another. Taking this approach will allow for the re-examination of the theoretical explanations outlined in this chapter of how these types of words are processed.

Chapter 3 – Linguistic characteristics and norming data from a set of Spanish-English identical and near-identical interlingual homographs.

Introduction

Chapters 1 and 2 identified interlingual homographs as being a useful tool by which to study bilingual language processing, because they share spellings but have distinct meanings between languages, and subsequently, can help us to understand how bilinguals store their languages (i.e., integrated or separate lexicons), and whether bilinguals selectively or non-selectively access their languages. This chapter introduces a novel set of 102 identical and near-identical Spanish – English interlingual homographs with their associated psycholinguistic properties for each language (i.e., length, word frequency, age of acquisition, imageability, length, bigram sum, and orthographic and phonological neighbourhoods). Moreover, it provides RTs and accuracy norming data from Spanish – English bilinguals conducting a LTD in their L2 English. The novelty of this chapter is that it introduces a more extensive set of Spanish - English interlingual homographs than those used in previous studies (e.g., Macizo, Bajo, & Martín, 2010), and it includes near-identical interlingual homographs that, to our knowledge, had not been explicitly considered in this language pair. This database will compile psycholinguistic properties of the stimuli in both English and Spanish and therefore it is intended as a tool to be used in future research.

Interlingual homographs and cognates have been the most important sources of stimulus materials in studies attempting to unravel the process of bilingual word recognition due to their lexical similarity between languages (Lemhöfer& Dijkstra, 2004). Interlingual homographs are particularly useful to study bilingual word recognition because they can help us to understand how bilinguals store and access their languages at the orthographic *and* semantic level; however, have not been researched in such depth compared to the cognate literature. The cognate facilitation effect is robust and well documented in a variety of languages and tasks such as written LDTs (e.g., Cristoffanini, Kirsner, & Milech, 1986, De Groot & Nas, 1991; Lemhöfer et al., 2008; Lemhöfer& Dijkstra, 2004; Peeters, Dijkstra, & Grainger, 2013; Poort & Rodd, 2017; Sáchez-Casas, García-Albea, & Davis, 1992; Van Hell & Dijkstra, 2002) and sentence paradigms (e.g., Duyck, van Assche, Drieghe, & Hartsuiker, 2007; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009; Van Hell & De Groot, 2008). Whereas, the processing of interlingual homographs generally comes at a cost (commonly coined as *inhibitory effects*): bilinguals take longer to process these words compared to matched controls in a variety of lexical decision and priming tasks (e.g., De Bruijn, Dijkstra, Chwilla, & Schriefers, 2001; De Groot, Delmaar, & Lupker, 2000; Dijkstra, De Bruijn, Schriefers, & Ten Brinke, 2000a; Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Dijkstra, Timmermans, & Schriefers, 2000b; Lemhöfer & Dijkstra, 2004; Libben & Titone, 2009; Macizo, Bajo, & Martín, 2010; Poort, Warren, & Rodd, 2016; Poort & Rodd; 2017; 2018; Titone, Libben, Mercier, Whitford, & Pivneva, 2011). Such effects are more strongly evident in tasks that probe the competing meanings of interlingual homographs. For instance, Macizo et al. (2010) used a semantic judgement relatedness task where participants determine whether pairs of words are semantically related to one another and found that Spanish - English bilinguals activated and then inhibited that non-target meaning of the interlingual homograph. Bilinguals responded more slowly to pairs of words involving homographs (e.g., PIE - TOE; pie meaning *foot* in Spanish) compared to pairs without homographs (e.g., LOG - TOE). In English the two pairs of words involving the homograph are unrelated (e.g., PIE – TOE), but when the Spanish meaning of the homograph is activated then there is a semantic relationship between the words. Thus, slower RTs to these pairs can be taken as evidence that the participants L1 has been activated.

Additionally, participants were slower when the pair included the translation of the Spanish meaning of the interlingual homograph (e.g., FOOT - HAND) after a trial involving an interlingual homograph (e.g., PIE - TOE), but not after a trial involving a control word. The results suggest that the non-target Spanish meaning of the homographs was activated initially but then inhibited as it was not relevant to the task.

However, it should be noted that some researchers have not found any interlingual homograph effects at all (e.g., Dijkstra et al., 2000b) or have found that these inhibitory effects are dependent on other variables such as the nature of the stimuli. For example, Dijkstra et al. (1998) conducted three LDTs which used Dutch – English interlingual homographs. In the first LDT, only English words and nonwords were included and the participants had to decide if a shown word was an English word or not. There were no differences in RTs for interlingual homographs compared to control words, suggesting no interference from the non-relevant L1. In the second LDT, L1 Dutch words were added to the stimulus material and the participants were instructed to treat the Dutch words as nonwords. In this case, bilinguals were slower to process interlingual homographs compared to controls. In the last LDT, the same materials were used but the task was different from the first two LDTs. Bilinguals were instructed to give a “yes” response when the word presented was English or Dutch and a “no” response when the word did not exist in either language. In this task, a facilitatory effect arose for interlingual homographs. This study highlights the differences that can arise when processing interlingual homographs depending on the context they are embedded, and the role of list composition. These findings are in contrast to the cognate literature that is robust; the interlingual homograph literature is still mixed in its findings.

Furthermore, in the cognate literature both identical and near-identical cognates (e.g., the words MAPA in Spanish and MAP in English) have been investigated, and what is found is that the overall magnitude of cognate facilitatory effects is moderated by differences in form overlap. Stronger facilitation effects are produced for identical cognates compared to near-identical cognates and the facilitation effect reduces in magnitude with decreasing form overlap (e.g., van Assche et al., 2011; Dijkstra et al., 2010; Peeters et al., 2013). Therefore, while investigating interlingual homographs is not a novel idea, studies have mainly focused on stimuli with identical spellings across languages. However, similarly to near-identical cognates, there are pairs of words that are similar across languages – not identical, but close – that could loosely be considered to share features with true interlingual homographs. These are referred to as near-identical interlingual homographs. For example, the English word *CARPET* refers to a type of flooring and the Spanish word *CARPETA* means *folder* in Spanish. The two words differ by only a single letter but have vastly different meanings in the two languages. Research specifically considering the processing of near-identical interlingual homographs is limited. Thus, an open question exists concerning whether near-identical interlingual homographs show the same effects as their identical counterparts. One study that has specifically explored orthographic overlap in interlingual homographs is by Di Betta, Okurowska, and Morgan (2015) who used both identical and near-identical interlingual homographs in a written sentence priming paradigm. Di Betta et al. (2015) asked proficient Polish-English bilinguals to read sentences biased towards the L2 English meaning of the interlingual homographs. The interlingual homograph was presented at the end of the sentence and acted as a prime for the following word (target) that could be related to L2 English meaning, to the L1 Polish meaning or was unrelated to either meaning. For instance, an example of a near-identical interlingual homograph is as follows: ‘Last summer we went camping in our CARAVAN’, where the last word looks like the Polish word KARAWAN meaning *hearse*. Overall, findings showed significant priming effects for both meanings of the interlingual homographs in the bilingual sample. However, when identical and near-identical interlingual homographs were considered separately, activation of the irrelevant L1 meaning was evident with near-identical homograph primes (facilitatory priming effects) but inhibition of the L1 was found with the identical interlingual homograph primes. These findings provide further evidence that all possible word candidates are activated regardless of language (Dijkstra & van Heuven, 2002), but also tentatively indicate that identical and near-identical interlingual homographs may be processed differently, at least in tasks that rely heavily on semantics.

As we have seen in Chapter 2, the Bilingual Interactive Plus model (BIA+; Dijkstra & van Heuven, 2002) suggests that interlingual homographs are represented by two competing representations which interfere with each other (Dijkstra & van Heuven, 2002; van Heuven & Dijkstra, 2010). The inhibition effects seen in Di Betta et al.'s (2015) work is explained by the BIA+ model as a strong effect of lateral inhibition, as, like with any other two words, the orthographic nodes of the interlingual homograph inhibit each other. However, because these representations are identical, this competition is stronger compared to non-ambiguous words which have one meaning, as both meanings of an interlingual homograph are activated until a relevant meaning is selected. Whereas, for the near-identical interlingual homographs the inhibition effect is reduced, so much so that having orthographic nodes that are not identical allows for the irrelevant meaning to be selected. Therefore, this would suggest that different control mechanisms in the lateral inhibition may be in place when processing identical compared to near-identical interlingual homographs. However, more research needs to be conducted to be able to reach firm conclusions about processing differences. Therefore, one novel aspect of this study is its aim to assess whether there are differences in how identical and near-identical interlingual homographs are processed in another language (Spanish), or whether the findings reported above are a) unique to Polish or b) unique to priming studies.

In order to answer these questions, this study focuses on exploring the processing of Spanish-English interlingual homographs at a linguistic level through a lexical decision task (LDT). In an LDT, participants make a decision about whether combinations of letters are words or not. For instance, when participants see the word “PLANT” they would respond “yes” because this is a real English word, but if participants saw the letters “SNISKY” then they would respond with “no” because this is not a real English word. The task was introduced in the 1970’s by Meyer and Schvaneveldt with the aim to investigate word recognition, semantic memory, and how we retrieve information from it (Meyer & Schvaneveldt, 1971; Schvaneveldt & Meyer, 1973; Meyer, Schvaneveldt, & Ruddy, 1975). While widely used, it is suggested that LDTs involve a high degree of noise (Diependaele, Brysbaert, & Neri, 2012), and therefore may not be best task for measuring lexical access. However, LDTs can be used to examine memory (Hicks, Franks, & Spitler, 2017) which is of particular relevance to this thesis as one of the questions in the literature introduced to you in chapter 2 is how bilinguals store and retrieve their languages (i.e., independent or integrated lexicons). Since the use of LDTs (in terms of analyses of accuracy and RTs) has produced inconsistent results in the literature when processing interlingual homographs (see above), the current study opted to assess the activation of the participants' L1 and L2 in the processing of the interlingual homographs in a more implicit manner. In other words, the interlingual homographs' linguistic properties in both Spanish and English were used as predictors of our bilinguals' performance in a L2 LDT. Furthermore, while LDTs are often combined with priming and have found that individuals respond to target words that were primed by a semantically related word faster compared to those target words primed by an unrelated word (e.g., individuals would be faster at responding to the word HOSPITAL when it was primed by the word DOCTOR, compared to the word BUTTER; Meyer & Schvaneveldt, 1971; Schvaneveldt & Meyer, 1973; Meyer et al., 1975); the current study did not include the element of priming because as stated above, it wanted a pure measure of homograph word recognition, hence to explore interlingual homograph processing in an implicit manner without the interference of other variables such as priming.

Linguistic properties

To further understand how language(s) are processed, researchers have examined whether specific linguistic properties of words influence the ways in which words are processed. It is well known that not all words are recognised with the same speed, and the literature has assessed the impact of a variety of linguistic properties on visual word recognition. However, in the case of interlingual homographs this is further complicated by the fact that the same orthographic form exists in two languages but the grammatical, pragmatic, and phonological rules governing usage means that the characteristics of a word will likely differ across languages. In what follows, we briefly highlight some of the key variables that have been shown to influence word recognition in bilingual and monolingual speakers that will be considered in this study.

Word and Bigram frequency.

Word frequency is considered to be one of the most important variables to control for in lexical tasks, because frequency effects are a common finding in visual word recognition in tasks which involve making a lexical decision (e.g., Connine, Mullennix, Shernoff, & Yelen, 1990; Frederiksen & Kroll, 1976; Gerhand & Barry, 1999; Hino & Lupker, 2000; Hudson & Bergman, 1985; Perea & Rosa, 2002; Richardson, 1976; Rubenstein, Garfield, & Millikan, 1970). Frequency effects refer to the fact that high-frequency words in a given language are processed faster in comparison to low-frequency words (Monsell, Doyle, & Haggard, 1989). The frequency effect is proposed to occur because high frequency words are easier to retrieve or be activated in the lexicon due to them being encountered and/or used more often than low frequency words (e.g., Brysbaert, Keuleers, & New, 2011; Frederiksen & Kroll, 1976; Hino & Lupker, 2000; Perea & Rosa, 2002). Another type of frequency that is important to look at is *bigram frequency* which is the frequency with which pairs of letters within words co-occur in a specific language. Bigram frequency is important to consider as it has been reported to be influential in tasks involving word recognition, with lower bigram frequency eliciting faster responses (e.g., Briederman, 1966; Broadbent & Gregory, 1968; Conrad, Carreiras, Tamm, & Jacobs, 2009; Massaro & Cohen, 1994; Owsowitz, 1963; Rice & Robinson, 1975; Rumelhart & Siple, 1974; Westbury & Buchanan, 2002). Several studies have reported significant bigram effects in lower-frequency words but not in high-frequency words and this considered to be because the word is more unique (for more information, see the review by Chetail, 2015; Gernsbacher, 1984). However, in contrast to the studies above, some researchers such as Andrews (1992) have suggested that bigram frequency has no effect on LDTs. Andrews (1992) manipulated bigram frequency and the size of the orthographic neighbourhoods (ONs). Orthographic neighbourhoods are frequently defined according to Coltheart Davelaar, Jonasson, and Besner’s (1977) as the *N* metric, which refers to the number of words that can be created by substituting one letter of the word to create orthographic neighbours (ONs). For instance, for the word 'coffee', one of its ONs would be 'toffee' (Marian, Bartolotti, Chabal, & Shook, 2012). ON is another linguistic property that is important in word recognition research as it can influence lexical access and is closely related to bigram frequency. As these two variables have been found to be highly correlated, Andrews (1992) aimed to see whether the effects of orthographic neighbourhood could be attributed to bigram frequency in monolingual speakers. An effect of orthographic neighbourhood was found, but there was no effect of bigram frequency on the RTs which would suggest this linguistic property does not affect word processing in an LDT, and that ON effects are not caused by bigram frequency. Furthermore, the British Lexicon Project (Keuleers, Lacey, Rastle, & Brysbaert, 2012) which is a database that contains lexical decision data for English words and non-words, replicates the findings of Andrews (1992), and also finds no effects of bigram frequency in a LDT. Therefore, strengthening the findings that bigram frequency is not influential in lexical decision making; although, the sample used was a monolingual sample and this thesis utilises bilinguals as the experimental group. Consequently, rather than questioning its impact, Schmalz and Mulatti (2017) state that bigram frequency should be mainly treated as a potential confounding variable, where researchers match for it during item selection as a conservative approach to creating a controlled stimulus set. Despite some studies findings no effects, as we have seen, other studies have (e.g., Conrad et al., 2009; Massaro & Cohen, 1994), and therefore, bigram frequency will be included as a variable of interest in the current chapter. Furthermore, bigram frequency has not been investigated in conjunction with interlingual homographs in an LDT, and therefore, it is unknown whether bigram frequency in the bilinguals' L1 and L2 has an impact on the processing of interlingual homographs. Lastly, by providing the bigram scores of each interlingual homograph it adds to the database which researchers can use to create a well-controlled stimulus set.

Word length.

The *length of a word* (how many letters it has) has also been found to influence performance in word recognition tasks. Generally, as word length increases so do the response times and the errors made (e.g., Balota, Cortese, Sergent-Marshall, Spieler, &

Yap, 2004; New, Ferrand, Pallier, & Brysbaert, 2006). For instance, New et al. (2006) focused on RTs, and took stimuli from the English Lexicon Project (a large data set of over 40,481 words; Balota et al., 2004) to further examine the influence of stimulus length on lexical decisions. The authors found that words between 3 to 5 letters long elicited faster responses than longer words, with each additional letter increasing response time.

Orthographic and phonological neighbourhood.

Orthographic and phonological neighbourhoods have also been found to affect word recognition, both within and between languages. Phonological neighbourhoods are calculated similarly to orthographic ones, but instead of depending on grapheme substitution, phonological neighbours (PNs) are created by substituting one phoneme of the word (Luce, Pisoni, & Goldinger, 1990). Generally, in LDTs, words with larger orthographic neighbourhoods are recognised more quickly in comparison to those with smaller neighbourhoods (e.g., Andrews, 1989; 1992; Forster & Shen, 1996; Sears, Hino, & Lupker, 1995). PNs work in a similar way to ONs in that words with many PNs are easier to recognise than words with fewer PNs, and produce faster RTs (e.g., Chen, Vaid, Boas, & Bortfeld, 2011; Yates, 2005; Yates, Locker, & Simpson, 2004). For instance, Yates et al. (2004) investigated phonological neighbourhoods and their role in the processing of words in two written LDTs. The first experiment revealed that words with larger phonological neighbourhoods were responded to faster and fewer errors were made in comparison to words with smaller phonological neighbourhoods. In the second experiment, Yates et al. (2004) replicated the effects of experiment 1 and found that increasing the number of PNs of a word, while holding the number of ONs constant, produced faster responses and fewer errors were made. What this study demonstrates is that PNs are important in visual LDTs, as they do affect participant's performance. Neighbourhood effects are thought to occur because they result in greater semantic activation (Holcomb, Grainger, & O'Rourke, 2002), which, in turn, aids individuals in tasks such as LDTs as the words with more neighbours are semantically more active in comparison to those with lower neighbourhoods. However, some researchers argue that LDTs do not always require semantic access (e.g., Grainger & Jacobs, 1996; Playfoot, Billington, & Tree, 2018; Poort & Rodd, 2019; Poort et al., 2016); in this case, instead of the meaning of the words being spread, it is the sound of words, creating a phonological spreading, instead of a semantic one.

Neighbourhood effects are also present between languages (e.g., Mulder, van Heuven, & Dijkstra, 2018; van Heuven et al., 1998). For instance, a study by van Heuven et al. (1998) used Dutch - English bilinguals to investigate how the recognition of target words belonging to the L2 (English) was affected by the existence of ONs from the same language or their L1 Dutch. Increasing the number of ONs in the L1 slowed RTs, while an increase in ONs in the L2 elicited faster responses. van Heuven et al.'s (1998) findings indicate that bilingual word recognition is influenced not only by the number of ONs in the target language, but also by the neighbourhood density in the non-target language. It is therefore important for researchers to be able to consider neighbourhood densities from both the target and non-target languages as this has processing implications for bilinguals.

Age of acquisition.

Age of acquisition (AoA; also known as order of acquisition) is also an important linguistic variable and refers to the moment in time when words are learned by individuals (e.g., Izura & Ellis, 2004; Izura et al., 2011, Stewart & Ellis, 2008). Words learned earlier in life are generally responded to faster and with better accuracy than words that have been acquired later, hence giving them a processing advantage. The AoA effect would suggest that earlier acquired words hold a sort of privileged position in the lexical or semantic network. This ‘network privilege’ is supported by studies showing that earlier acquired words have a greater resistance to brain damage (e.g., Johnston & Barry, 2006; Juhasz, 2005), and are influential in predicting what information is more likely to be lost after brain injury and in conditions such as aphasia and Alzheimer’s disease (e.g., see Brysbaert & Ellis, 2016 for a review). AoA effects can be explained by the semantic growth model (Steyvers & Tenenbaum, 2005) that proposes that these effects are due to enhanced semantic connectivity that reflects a rich underlying semantic network in the individual. Therefore, words that are acquired earlier have made more connections compared to those learned later and are subsequently, connected better (Brysbaert, van Wijinendaele, & de Deyne, 2000), and are less likely to be lost. Likewise, the network plasticity hypothesis proposes that AoA effects are due to the decrease of network plasticity during life development (Ellis & Lambon Ralph, 2000). AoA is implicated in learning mechanisms more generally (e.g., Catling, Dent, Preece, & Johnston, 2013; Stewart & Ellis, 2008), as the effects of AoA commonly appear when stimuli have to be learned over a period of time in an accumulative manner, which further demonstrates that AoA may play a key role in a person's learning (Johnston & Barry, 2006).

AoA effects are present in both lexical and non-lexical tasks (e.g., Bonin, Chalard, Méot, & Barry, 2006; Brysbaert & Cortese, 2011; Brysbaert et al., 2000; Catling, Dent, & Williamson, 2008; Cortese & Khanna, 2007; Holmes, Fitch, & Ellis, 2006; Juhasz, 2005; Monaghan & Ellis, 2002; 2010; Richards & Ellis, 2009); behavioural studies, ERP components, and eye fixation durations (e.g., Cuetos, Barbón, Urrutia, & Domínguez, 2009; Ellis, Burani, Izura, Bromiley, & Venneri, 2006; Fiebach, Friederici, Müller, von Cramon, & Hernandez, 2003; Juhasz & Rayner, 2006; Morrison

& Ellis, 1995; Pérez, 2007; Weekes, Chan, & Tan, 2008). Moreover, AoA effects are not limited to one language, effects have been found in many languages and are present in both bilingual and monolingual speakers (e.g., Assink, van Well, & Knuijt, 2003; Hirsh, Morrison, Gaset, & Carnicer, 2003; Izura & Ellis, 2004). Indeed, AoA effects that mirror those in monolinguals have been found in those who learned a L2 after childhood (e.g., Alija Fernández & Cuetos Vega, 2006; Ferrand et al., 2011; Izura & Ellis, 2002; Liu, Hao, Shu, Tan, & Weekes, 2008; Menenti & Burani, 2007; Raman, 2006; Wilson, Cuetos, Davies, & Burani, 2013; Wilson, Ellis, & Burani, 2012). This research demonstrates that there is strong evidence of the robustness of AoA effects in a variety of languages and task paradigms.

Imageability.

Faster and more accurate responses are also associated with words that easily produce a mental image or sensory experience (e.g., Balota et al., 2004; Cortese & Schock, 2013; Cortese, McCarty, & Schock, 2015), and this is commonly referred to as the *imageability effect* (Paivio, Yuille, & Madigan, 1968). For example, in Gilhooly and Logie's (1980) norms the words "apple", "tent", and "teacher" generated a high-imageability score, whereas the words "recreant" and "fane" were rated as lower in imageability. Previous research has shown that highly imageable words are recognised and memorised better in a variety of tasks such as lexical decision (e.g., Balota et al., 2004; Cortese & Schock, 2013; Kounios & Holcomb, 1994; Kroll & Merves, 1986; Schwanenflugel, Harnishfeger, & Stowe, 1988), semantic similarity judgment tasks (e.g., Sabsevitz, Medler, Seidenberg, & Binder, 2005), free recall of word lists (i.e., list learning; e.g., Kennet, McGuire, Willis, & Warner Schaie, 2000; Mellet, Tzourio, Denis, & Mazoyer, 1998; Paivio, 1967), cued recall (e.g., Paivio, 1963, 1965; Yarmey & O'Neill, 1969), word recognition memory (e.g., Cortese, Khanna, & Hacker, 2010;

Cortese et al., 2015), and in reading out loud and word production paradigms (e.g., Alario et al., 2004; Balota et al., 2004; Bleasdale, 1987; Coltheart, Laxon, & Keating, 1988; Strain, Patterson, & Seidenberg, 1995).

Higher imageability is associated with easier access to an individual's semantic representations, indeed, the effect has been attributed to different types of representations in the mental lexicon for high and low imageability words (e.g., Paivio, 2014). For instance, the dual-code hypothesis proposes that words have two potential codes of representation, a visual and a verbal code (Paivio 1971; 2013; Clark & Paivio, 1991). High-imageability words have an advantage over low-imageability words because they possess both visual *and* verbal representations, and not just verbal representations. Effects have been found in many languages and in L1 and L2 learned after childhood; though, tasks exploring imageability have mainly used aphasic bilingual patients (e.g., Kiran & Tuchtenhagen, 2005; Mortensen, Berntsen, & Bohn, 2015; Poncelet, Majerus, Raman, Warginaire, & Weekes, 2007). Kiran and Tuchtenhagen (2005), in addition to bilingual aphasic individuals, also looked at the performance of healthy English - Spanish bilinguals (who learnt both languages growing up but described themselves as more dominant in English) in a naming and semantic priming task. The authors found that across tasks and languages, performance was better in the bilinguals' dominant English compared to Spanish, and that responses were faster and more accurate for concrete words compared to abstract words. Concrete words (e.g., apple) are easier to picture in your mind and thus have higher imageability scores compared to abstract words (e.g., truth) that are harder to imagine. This study is particularly interesting, as it shows that imageability effects were present in the both the bilinguals' L1 and L2 which could suggest that semantic information is represented in an integrated manner in the bilinguals' lexicon and that they have non-selective access over their languages (Dijkstra & van Heuven, 2002).

Overall, the above findings have shown that the manipulation of a number of linguistic factors can be valuable in exploring the organization and function of the language processing system in both bilingual and monolingual individuals. Though, as the list of linguistic properties known to affect performance gradually increases, experimentation becomes more difficult if one is to avoid confounding them with the particular property of experimental interest. Furthermore, as discussed earlier, interlingual homographs are being used more in research because they are influential in the investigation of word recognition in the bilingual lexicon. However, specific characteristics of individual words that influence processing have not been systematically studied in Spanish – English interlingual homograph experiments, specifically utilising both identical and near-identical interlingual homographs. Consequently, accessible lists of identical and near-identical interlingual homographs and norming data are not readily available for researchers to easily compare between languages. Therefore, these challenges can be dealt with by making use of standardised stimuli sets and databases where researchers can gather linguistic property information on the set of experimental words and allow for the manipulation or control of these linguistic characteristics.

The present experiment: Study 1

All the studies outlined in this chapter received ethical approval from Sheffield Hallam’s University Research Ethics Committee (ER5583251).

This study has two aims. The first aim is addressed in Experiment 1a and it was to create a database and collate a list of both identical and near-identical Spanish – English interlingual homographs and to describe them in terms of key lexical, sub lexical, and semantic characteristics. However, given the novel nature of near-identical interlingual homographs, the ambiguity of these words was first validated in a survey with Native Spanish speakers who were also proficient in the English language. Moreover, while many rating norms for imageability and AoA already exist in many languages, fewer are available for both meanings of interlingual homographs, making cross-language comparison a challenge. Therefore, a survey was conducted to collect imageability and AoA ratings in English *and* Spanish for both identical and near-identical interlingual homographs. These two linguistic characteristics are important to collect as discussed above, there is evidence to suggest that they can influence the processing of words (e.g., Assink et al., 2003; Balota et al., 2004; Cortese et al., 2015; Izura & Ellis, 2004).

Subsequently, a by-product of this chapter is that there will be an easily-accessible list of both identical and near-identical Spanish – English interlingual homographs with varying degrees of spelling similarity, with their associated linguistic characteristics in both languages and to be used in future research. The database will enable researchers to match stimuli more rigorously in future experiments. In order to appropriately design studies using interlingual homographs as stimuli, being able to manipulate or control these linguistic properties is vital.

Given that a variety of linguistic properties have been shown to affect word recognition in L1 and in L2 the <u>second aim</u> is to investigate the impact orthographic overlap and linguistic properties have on bilinguals' performance in a LDT in their L2 English. The LDT was chosen as it allows for the recording of accuracy and RTs and is ideal as one can manipulate the language the bilingual participants complete the task in, and subsequently can measure whether there is interference from the alternate language's linguistic characteristics. A visual written LDT was chosen instead of an auditory LDT because it removed the confounding variable of accents, since previous research has found that familiarity with the accent of the speaker has a facilitatory effect on lexical processing (Adank, Evans, Stuart-Smith, & Scott, 2009; Clopper, 2021; Floccia, Goslin, Girard, & Konopczynski, 2006).

The study is novel as there has been no studies that has exclusively investigated identical and near-identical interlingual homographs at a lexical level, and their associated linguistic characteristics. Given that the linguistic properties outlined above have been shown to affect lexical and semantic processing of words in monolinguals and bilinguals operating in their L2, it is worth examining how these variables also affect the processing of words which share the same (or similar) form in both languages. This constitutes the most comprehensive assessment of the impact of psycholinguistic variables on Spanish-English interlingual homograph processing to date.

Based on previous research where frequency effects are common in LDTs (e.g., Connine et al., 1990; Perea & Rosa, 2002), it is predicted that high-frequency L2 words will be processed more quickly in comparison to low-frequency words as the task is in the bilinguals' L2. However, if the L1 is activated then the opposite might be true for the Spanish meaning, hence instead of having a facilitatory effect, high-frequency L1 Spanish words may interfere and slow down lexical decision-making time as the task is in the bilinguals' L2 English and not the L1. It is difficult to make a prediction for bigram frequency due to the unreliable nature of the linguistic property in word recognition; some researchers have found a bigram frequency effect in an LDT (e.g., Westbury & Buchanan, 2002), and others have not (e.g., Keuleers et al., 2012). Therefore, it is unknown whether bigram frequency will have an effect in this study when processing interlingual homographs. Based on previous research (e.g., Balota et al., 2004; New et al., 2006), it is predicted that as word length increases, so will the response times and errors. Moreover, based on previous research (e.g., Sears et al., 1995; Yates et al., 2004) it is predicted that as the L2 orthographic and phonological neighbourhoods increase, the faster responses will be. However, similar to what was predicted for frequency, as neighbourhoods of the bilinguals' non-target language increase, the activation of the bilinguals' non-target language appears to increase too and has an inhibitory effect on the bilinguals' lexical decision performance (Mulder et al., 2018). Therefore, it is predicted that as the orthographic and phonological neighbourhoods of the L1 increase, the slower the response time will be when making lexical decisions. Additionally, it is predicted that words learned earlier in life will be responded to faster and with fewer errors compared to words acquired later in life (e.g., Izura et al., 2011, Stewart & Ellis, 2008). Lastly, based on previous research that has found that faster and more accurate responses are associated with words that can produce a mental image or sensory experience (e.g., Cortese & Schock, 2013; Cortese et al., 2015), it is expected that those L2 meanings that are more imageable with lead to faster response times and fewer errors compared to those meanings that are difficult to produce a mental image or sensory experience. Though, if the L1 meaning of the interlingual homograph can produce a mental image or sensory experience and the L1 imageability is activated, then this will lead to slower response times and more errors due to interference from the non-target language.

Study 1a: Creating the database

The stimuli consisted of a 102-word list formed of 50 identical and 52 near-identical interlingual homographs. All interlingual homographs were manually selected from two sources: The Oxford Spanish Dictionary (Galimberti Jarman, Russell, Styles, Parker, & Huelskamp, 1994) and The Collins Spanish-English dictionary (Smith, Bermejo Marcos, & Chang-Rodríguez, 1986). For each interlingual homograph, indices of frequency of occurrence (word frequency), lexical similarity (i.e., bigram frequency, neighbourhood size, and spelling difference in near-identical interlingual homographs), imageability, and AoA (the age the words were learnt) were collated for English and Spanish forms. A complete list of the stimuli, along with their linguistic characteristics, can be found in Appendix 1.

Word Frequency.

The word frequency measures presented here were taken from SUBTLEX-UK (van Heuven, Mandera, Keuleers & Brysbaert, 2014) for the English words and from SUBTLEX-ESP (Cuetos et al., 2012) for the Spanish words. Both corpora are based on language samples from film subtitles (201.3 million words in the case of SUBTLEX-UK and 41.5 million words in the case of SUBTLEX-ESP). Results from more recent studies suggest that word frequency estimates based on television and film subtitles are better in predicting performance in word recognition experiments compared with traditional word frequency estimates obtained from books and newspapers (e.g., New, Brysbaert, Veronis, & Pallier, 2007; Cuetos, Nosti, Gutiérrez, & Brysbaert, 2011; Brysbaert, Keuleers & New, 2011). For example, Brysbaert and New (2009) showed that frequency estimations account for a high proportion of variance in naming times compared to more traditional frequency values such those reported in Kučera and Francis (1967), CELEX (Baayen, Piepenbrock, & van Rijn, 1993), HAL (Lund & Burgess, 1996), and Zeno (Zeno, Ivens, Millard, & Duvvuri, 1995). However, it should be noted that written texts may not reflect the language used by individuals in daily life, as writers use a more educated and refined register and this can lead to an under estimation of many common words, and an over estimation of words rarely used in everyday life (Cuetos et al., 2011). Written texts also tend to focus on lexical variation in order to avoid word repetition, which does not occur in spoken language (Cuetos et al., 2011; Brysbaert et al., 2011).

To standardise the frequency scores, the Zipf scale was used for both languages (Zipf, 1936; 1949). The Zipf is a logarithmic scale and is a log transformation of the frequency per million of words (fpmw) measure (i.e., log10 (fpmw) + 3); it roughly ranges from 1 (1-3 low frequency words) to 6 (4-6 high frequency words) or 7 (function words, pronouns, and verb forms like "have"). The scale makes it easier to compare across languages in contrast to other frequency measures, because the use of the Zipf scale to measure frequency in both languages enables an index of the same interpretation across all stimuli used, such as the corpus of words analysed (van Heuven et al., 2014; Cuetos et al., 2011). This scale is also implemented in models of the bilingual lexicon (e.g., Dijkstra et al., 2019).

Estimates of lexical similarity.

Interlingual homographs have been described here in terms of their similarity with other words in English and Spanish. Neighbourhood size (both orthographic and phonological) and Levenshtein Distance were considered as measures of similarity at the word level, and bigram frequency as an index of the prevalence of letter-level patterns in the target languages. Orthographic and phonological neighbourhood sizes were taken from 'Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities' (CLEARPOND; Marian et al., 2012) for both English and Spanish words. Orthographic and phonological neighbours were defined as words that could be formed by changing only one phoneme or letter of the target word. Bigram sums were also taken from CLEARPOND (Marian et al., 2012). Identical homographs generated one bigram sum score in the CLEARPOND as the exact same spelling was used. However, for the near-identical homographs, two bigrams scores were generated: one for the Spanish spelling and the other for English spelling.

To calculate how different the spellings of two near-identical interlingual homographs are between languages, the Levenshtein Distance was calculated. This is the orthographic difference between two words, that is, the minimum number of insertions, deletions, or substitutions needed to change one word into the other (Levenshtein, 1966). The Levenshtein Distance mean distance for all 52 near-identical interlingual homographs was 1.62 (*SD* = 1.03). Most of the words had a score of 1 (*n* = 34), followed by 2 (*n* = 9), 3 (*n* = 6), 4 (*n* = 1), with the maximum Levenshtein Distance being 5 (*n* = 2).

Validation Survey: Confusability ratings of the near-identical interlingual homographs

The aim of this study was to validate the Spanish – English near-identical interlingual homographs by obtaining a measure of how confusable native Spanish speakers found the English and Spanish orthographic forms to be with another.

A total of 29 native Spanish speakers who were fluent in English were recruited from Prolific (https://prolific.ac/, 2020) and the researcher’s social media sites. To be included in the study the following criteria were applied: no known language impairments, native Spanish speaker, and fluent in English and Spanish only. There were 14 males, and 15 females, with a *mean* age of 29 years (*SD* = 10.6; range = 18 –56 years). Participants also completed the 'Lexical Test for Advanced Learners of English' (LexTALE; Lemhöfer & Broersma, 2012) and the Spanish equivalent (LexTALE-ESP; Izura, Cuetos & Brysbaert, 2014) to ensure high-proficiency in both languages. Participants scored 84.8% (*SD* = 4.9) in the LexTALE, and 95.9% (*SD* = 2.9) in the

LexTALE-ESP signalling high-proficiency in both languages. All were educated to a minimum of A-level or equivalent standard. A total of 16 were students, and 13 were in full-time work. The survey was in Spanish, and bilinguals were asked to rate pairs of words, one in Spanish and the other in English (e.g., compromiso – compromise), which are near-identical interlingual homographs, on 3-point scale (1 = never confuse, 2 = sometimes confuse, 3 = always confuse). Bilinguals were asked to focus on the written content, and not the sound of the words. For each near-identical interlingual homograph at least 25% of participants rated that they confused the English and Spanish words either “sometimes confuse” or “all of the time” (*mean* = 1.7, *SD* = 0.2). While these findings provide tentative support that bilinguals in this study did find the English and Spanish orthographic form to be confusable with one another, there were near-identical interlingual homographs that were rated more confusable than others. For instance, TRAMP-TRAMPA (the latter meaning *trap* in Spanish) scored 2.12 whereas, PARTY – PARTO (the latter meaning *birth* in Spanish) was scored the lowest with 1.45. Subsequently, while the results show that at least 25% of the stimuli were considered ambiguous, these results should be taken with caution as the confusability is difficult to interpret given the small range of scores in the Likert scale. Future research should look to increase the Likert scale to allow for clearer comparison between pairs and could contain either five or seven response categories (Bearden, Netemeyer, & Mobley, 1993; Peter, 1979; Shaw & Wright, 1967), instead of 3. A full list of the materials and findings can be found in Appendix 2 – 4.

Imageability and AoA ratings Survey in English and Spanish

As mentioned in the introduction, norms for imageability and AoA already exist in many languages, but far fewer are available for *both* meanings of interlingual homographs in English and Spanish, making cross-language comparison a challenge.

Therefore, the aim of these surveys was to obtain subjective imageability and AoA ratings in English and Spanish for the 102 interlingual homographs to be added to the database.

One of the ways researchers have collected AoA estimates is by asking groups of people to estimate the age at which they believe they learned a list of words. Subjective measures such as AoA estimates are particularly useful in bilingualism, as words from a L2 may have been learnt at a different time from a L1 such as in a school environment. AoA estimates have been shown to correlate highly with objective measures (e.g., reading lists) of AoA demonstrating reliability (e.g., Carroll & White, 1973; Gilhooly & Gilhooly, 1980; Pérez, 2007). Overall, as discussed earlier the facilitation effect in early acquired words is robust and is a particularly important variable to consider in semantic tasks, because AoA influences the ease of access of individuals' semantic representations. Similarly, to AoA, imageability ratings have also been collected, studied, and used in various psycholinguistic studies (see Rofes et al., 2018 for review). Imageability estimates are also useful in bilingualism research, as the linguistic variables form part of the richness of a semantic representation of a word (Breedin, Saffran, & Coslett, 1994; Plaut & Schallice, 1993). In this study, two surveys were conducted, one on native English speakers (NEs) in order to collect AoA and imageability ratings in the English language, and the other on monolingual Spanish speakers to collect the same type of ratings in Spanish.

Method

Participants [English survey]

A total of 90 native English speakers (20 males and 70 females), whose ages ranged from 18 to 62 years (*mean* = 25 years, *SD* = 10.6), all with British nationality and living in the United Kingdom, with no experience of the Spanish language, took part in the English imageability and AoA rating survey. A total of 37 had no experience with another language apart from English, 53 had experience with another language (e.g., GCSE) but did not rate their proficiency above 3 (acceptable). A total of 4 participants reported a language impairment (e.g., dyslexia) and 4 preferred not to say. Participants from Sheffield Hallam University were offered course credits for their participation (*N* = 18). The sample consisted of 70 students and 20 individuals in full-time work.

Participants [Spanish survey]

A total of 55 native Spanish speakers (36 males and 19 females), aged between 18 and 62 years (*mean* = 30 years, *SD* = 8.6), with Spanish nationality, all living in Spain, and with no or limited experience of the English language, took part in the Spanish survey. A total of 16 had no experience with another language apart from Spanish, 39 had experience with another language but did not rate their proficiency above 3 (acceptable). No participants reported a language impairment. Participants were recruited using Prolific (https://prolific.ac/, 2019) and were offered the standard rate for their participation. Prolific is an online website that allows you to launch your online survey to find research participants that are registered and meet your experimental criteria. The criterion for the native Spanish speakers was that they had to be native speakers of the Spanish language, have limited or no prior knowledge of the English language, and be above the age of 18. The sample consisted of 21 students, 32 individuals in full time work, and 2 that were currently unemployed.

Design and Materials.

This was a survey design with the variables of interest being AoA and imageability. The experiment was set up in Qualtrics (Qualtrics, 2018) which is an online platform used to conduct surveys. A copy of the materials including the surveys in English and Spanish can be found under the supplementary material section (see Appendix 5 – 6). Furthermore, a correlational design was also used to explore the relationship of the imageability and AoA ratings with the rest of the linguistic variables collated (i.e., word and bigram frequency, word length, orthographic and phonological neighbourhood).

The imageability instructions and scale were adapted from Paivio et al. (1968) and required participants to indicate the ease with which each of the words evoked a mental image. Numbers in the scale were labelled and informed the participants of the different degrees of image-evoking difficulty. The Likert scale ranged from 1 (image aroused after long delay/not at all) to 7 (image aroused immediately) and participants dragged a slider to the appropriate number for each word. The AoA section of the survey required participants to estimate when they first had learned each of the words in the lists by typing the estimated age in a box located beside each word. This method has been used successfully in the past (e.g., Ghyselinck, De Moor & Brysbaert, 2000; Izura, Hernandez-Muñoz & Ellis, 2005; Izura & Playfoot, 2012; Montefinese et al., 2019).

The same imageability and AoA questionnaires described above were translated into Spanish and checked for language consistency by an academic who works in the United Kingdom but is a native Spanish speaker. The stimuli were altered slightly for the Spanish questionnaire in that the Spanish spelling of the near-identical interlingual homographs was used instead of the English. For example, instead of asking Spanish participants to rate the word NUDE (not a Spanish word), NUDO was rated (meaning *knot* in Spanish).

Procedure.

Participants firstly filled out a language background questionnaire, followed by the imageability and AoA ratings. Half the participants rated the imageability first and then the AoA, and vice versa. The surveys took approximately 15 minutes to complete.

Results

Collecting linguistic information for each language of the interlingual homograph, allows for a deeper exploration of these words' linguistic properties, and the impact of cross-language orthographic overlap. Note that a near-identical interlingual homograph has two spellings, one in English (e.g., NUDE) and one in Spanish (e.g., NUDO). Therefore, there are bigram frequency values for each of them in their respective languages in the database.

Descriptive statistics for each of the linguistic variables considered in this study are shown in Table 1 split between the orthographic nature of the interlingual homograph: identical or near-identical.

Table 1.

Descriptive statistics for each linguistic variable considered in the study.

		Mean	*SD*	*Min*	*Max*
AoA (years)					
	Identical	7.55	2.99	2.09	14.64
English	Near-identical	6.84	2.32	3.43	12.76
Spanish	Identical	6.77	1.69	3.76	11.30

	Near-identical	6.77	1.85	3.91	13.36
Imageability					
English	Identical	4.42	1.25	2.22	6.61
	Near-identical	5.18	1.29	2.46	8.23
Spanish	Identical	4.62	1.03	3.11	6.60
	Near-identical	5.54	.96	2.78	6.84
Bigram Frequency Sum					
English Bigram Sum	Identical	.04	.03	.00	.19
	Near-identical	.05	.05	.01	.22
Bigram Sum with English spelling (Spanish bigram sum)	Identical	.05	.03	.01	.14
	Near-identical	.05	.04	.01	.18
Bigram Sum with Spanish spelling (Spanish bigram sum)	Identical	.05	.03	.01	.14
	Near-identical	.06	.05	.01	.22
Length (number of letters)					
	Identical	4.5	1.47	2	10
	Near-identical	5.85	2.12	4	11
Spanish spelling	Near-identical	5.88	2.05	4	10
Orthographic Neighbourhood					
English	Identical	9.62	7.84	0	26
	Near-identical	4.85	5.58	0	23

Spanish	Identical	4.04	3.00	0	11
	Near-identical	4.48	4.75	0	15
Phonological Neighbourhood					
English	Identical	19.26	14.41	0	11
	Near-identical	11.35	12.37	0	40
Spanish	Identical	4.46	3.28	0	11
	Near-identical	5.13	5.73	0	17
Frequency					
English	Identical	4.08	.98	1.97	6.61
	Near-identical	4.03	.80	1.74	5.54
Spanish	Identical	4.24	1.08	1.98	6.86
	Near-identical	4.21	.60	2.72	5.41
Levenshtein Distance	Near-identical	1.62	1.03	1	5

Note: SD = Standard Deviation, *Min* = Minimum, *Max* = Maximum. The brackets represent the language, either English or Spanish.

A series of t-tests were conducted to compare the linguistic characteristics between languages (see Appendix 7). There were significant differences between English and Spanish for 3 variables: orthographic and phonological neighbourhoods, and bigram frequency. There were significantly more English ONs (*mean* = 7.19) compared to Spanish ONs (*mean* = 4.26; *t* (101) = 4.90, *p* < .001). Additionally, there were significantly more English PNs (*mean* = 15.23) compared to Spanish PNs (*mean* = 4.80; *t* (101) = 4.90, *p* < .001). Bigram frequency sum was significantly higher in Spanish (*mean* = .054, *SD* = .04) compared to English bigram scores (*mean* = .050, *SD*

= .04; *t* (101) = -2.80, *p* = .006). There were no significant differences between length, frequency, imageability, and AoA scores. Moreover, correlations for linguistic properties for the interlingual homographs can be seen in Table 2 (see Appendix 8).

Table 2.

Correlations between all the linguistic variables collated for the word list.

		English language						*Spanish language*							Near-identical only	
		I	**Bigram**	**Length**	**ON**	**PN**	**F**	***AoA***	***I***	***Bigram***	***Length***	***ON***	***PN***	***F***	**Bigram English Spelling**	**LD**
English language	**AoA**	-.087	**2.62***	**.314***	-.161	-.274*	-.620**	.169	-.184	**.242***	**.305***	-.178	-.179	-.150	**.254***	-.014
	I	-	-.129	-.017	-.056	-.070	-.023	-.087	.178	-.116	-.039	.105	.099	.009	-.150	**.195***
	Bigram		-	**.792****	-.309*	-.404**	-.319*	**.262***	-.129	**.896****	**.780****	-.324*	-.342**	-.212*	**.889****	**.438****
	Length			-	-.641**	-.688**	-.393**	**.314***	-.098	**.769****	**.987****	-.520**	-.524**	-.243*	**.737****	**.593****
	ON				-	**.874****	**.299***	-.161	-.161	-.339**	-.656**	**.541****	**.515****	**.270***	-.276*	-.374**
	PN					-	**.378****	-.274*	-.098	-.439**	-.699**	**.495****	**.475****	**.230***	-.394**	-.352**
	F						-	-.620**	-.023	-.257*	-.386**	**.328***	**.323***	**.202***	-.217*	-.134
Spanish language	***AoA***							-	-.560**	**.227***	**.381****	-.295*	-328*	-.366**	**.218***	.111
	I								-	-.155	-.104	**.210***	**.202***	.029	-.142	.157
	Bigram									-	**.768****	-.303*	-.355*	-.128	**.961****	**.451****
	Length										-	-.547**	-.556**	-.246*	**.726****	**.569****

	ON	-	**.929****	.119	-.245*	-.158
	PN		-	.158	-.281*	-.144
	F			-	-.113	-.049
Near-identical only	**Bigram English Spelling**				-	**.341****

Note: Significant positive correlations are reported in bold. Spanish linguistic variables are reported in italics. Correlations whose *p* was > .05 are reported as non-significant (NS); *p < .05; **p < .001; AoA = Age of Acquisition; I = Imageability; ON = orthographic neighbourhood; PN = phonological neighbourhood; F = Frequency; LD = Levenshtein Distance.

The correlations reported in Table 2 demonstrate that there were significant correlations between nearly every pair of variables. These will not be discussed exhaustively but will highlight key patterns. AoA and imageability are discussed in detail as these were the norms collated from the survey. Moreover, some of the stronger correlations are not of theoretical interest, particularly those related to the orthographic length of an interlingual homograph. For instance, there are strong positive correlations between word length in English and Spanish, and bigram frequency in both languages. The nature of the current stimuli list is such that the items are identical (or nearly) in Spanish and English, hence they usually have the same number of letters. Summed bigram frequency measures were used, so the longer the string, the greater the number of contributing bigrams, and the higher the overall bigram frequency. Furthermore, there were moderate negative correlations between length and orthographic neighbourhood size (i.e., as length goes up, neighbourhood size goes down) as has been observed in previous studies (e.g., New et al., 2006). The correlations between orthographic and phonological neighbourhoods were strong in both languages, with words that have a greater number of ONs also tending to have a greater number of PNs. This matches previous findings (e.g., Mulatti, Besner, & Job, 2003). Importantly there were also moderate positive correlations between orthographic and phonological neighbourhood size *across* English and Spanish. Given that it has been shown that orthographic neighbourhood effects can be eradicated when phonological neighbourhood is controlled (Mulatti et al., 2006) these correlations demonstrate that it would be important to consider cross-language neighbourhoods carefully in research when using bilingual samples.

Linguistic properties that show a significant *positive correlation* with English AoA ratings are: bigram sum, bigram sum with English and Spanish spellings, and length in both languages. Those with a significant *negative correlation* with English AoA ratings are: English PN and English frequency. These patterns indicate that interlingual homographs that have higher AoA ratings in English have higher bigram sum, bigram sum with English spelling, bigram sum with Spanish spelling, longer lengths in both languages; and those AoA ratings in English with higher scores have lower scores of imageability in English, and smaller English phonological neighbourhoods. Variables that were found to not be significant for English AoA were: AoA in Spanish, imageability in both languages, ON in both languages, Spanish PN, Spanish frequency, and the Levenshtein distance.

Linguistic properties that show a significant *positive correlation* with Spanish AoA ratings are: English and Spanish bigram sum, bigram sum with English spelling, and length in both languages. Those with a significant *negative correlation* with Spanish AoA ratings are: PN and frequency in both languages, Spanish imageability and ON. These patterns indicate that interlingual homographs that have higher AoA ratings in Spanish have higher bigram sum in both languages, bigram sum with English spelling, and longer lengths in both languages; and those AoA ratings in Spanish with higher scores have lower phonological neighbourhoods and lower scores in frequency in both languages, and Spanish imageability and orthographic neighbourhood. Variables that were found to not be significant for Spanish AoA were: AoA, imageability, and ON in English; and the Levenshtein distance.

Linguistic properties that show a significant *positive correlation* with English imageability ratings are: the Levenshtein distance. There were *no significant negative correlations* with English imageability ratings. These patterns indicate that interlingual homographs that have higher imageability in English have higher Levenshtein distance scores. Variables that were found to not be significant for English imageability were: English and Spanish bigram sum, bigram sum with English spelling, orthographic and phonological neighbourhoods in both languages, English and Spanish frequency and length, and Spanish imageability.

Linguistic properties that show a significant *positive correlation* with Spanish imageability ratings are: Spanish ON and PN. There were *no significant negative correlations* with Spanish imageability ratings. These patterns indicate that interlingual homographs that have higher imageability in Spanish have larger orthographic and phonological neighbourhoods. Variables that were found to not be significant for Spanish imageability were: English and Spanish bigram sum, bigram sum with English spelling, orthographic and phonological neighbourhoods in English, English and Spanish frequency and length, English imageability, and the Levenshtein distance.

While the list is limited in terms of how many interlingual homographs are shared between the English and Spanish language, this database is a valuable resource for researchers who want to use interlingual homographs as stimuli and to maintain appropriate linguistic control in their experiments. To summarise this section, Study 1a collated linguistic characteristics of a set of interlingual homographs and obtained ratings for both AoA and imageability in English and Spanish that will be used in Study 1b for the LDT. The next section will use the information collated in this study to investigate how these variables impact bilinguals' performance in a LDT.

Study 1b: Using the database - Lexical decision task and the role of L1 and L2 variables

Having collected information about the lexical characteristics of a set of Spanish-English interlingual homographs, the next aim was to determine the influence of these on written word recognition. Given that these linguistic properties have been shown to affect lexical and semantic processing of words by bilinguals performing in their second language, we were able to predict that response times would be faster (and accuracy better) for words with shorter lengths (e.g., New, 2006), higher in frequency (e.g., Brysbaert et al., 2011), higher in imageability (e.g., Paivio et al., 1968), and with a greater number of ONs and PNs in the bilinguals' L2 English (e.g., Chen et al., 2011; Yates et al., 2004). It was also expected that response times would be slower and accuracy lower for words with later AoA ratings (e.g., Ellis & Lambon Ralph, 2000). Furthermore, if languages are integrated and selected non-selectively (e.g., Dijkstra & van Heuven, 2002), then an effect of the L1 Spanish linguistic properties should be observed on accuracy and/or latency scores for the L2 English target words. If no influence of the L1 Spanish linguistic variables is observed, then it would support the notion that bilinguals are able to enforce some level of control over their languages (e.g., Green, 1998).

The research on Spanish – English near-identical homographs is limited to date, and there has not been a study that has directly looked at the effects of varying orthographic overlap in a bilingual's performance. Subsequently, there is no directly related previous research on which to base predictions. However, based on the cognate literature where facilitation effects are reduced as orthographic overlap is increased (e.g., van Assche et al., 2011; Dijkstra et al., 2010; Peeters et al., 2013), it is expected that there will be stronger L1 effects when bilinguals process identical interlingual homographs compared to near-identical interlingual homographs, which are real words only in the bilinguals' L2 English.

Method

Participants.

Thirty Spanish-English bilinguals (21 females and 9 males; *mean age* = 25.6 years, *SD* = 6.08) from the areas around Sheffield and Swansea took part in the study. No participants reported having a language impairment (e.g., dyslexia) or neurological disorder. Participants were given a £5 shopping voucher for their participation. All participants were asked to complete a language background questionnaire which asked participants about any previous languages they tried to learn and were asked to rate how well they could communicate with the English and Spanish languages in reading, writing, speaking and understanding on a 5-point Likert scale (0 (not at all), 1 (not well), 2 (I can, but with a lot of difficulty), 3 (well, but with a little difficulty), and 4 (very well). Participants self-reported their communication in English to be “very well” with a mean of 3.8 (*SD* = .7) and all participants rated their Spanish as 4. Bilinguals reported being dominant in their native language Spanish, though proficient and daily users of the English language. In addition to subjective measures, participants completed LexTALE (Lemhöfer & Broersma, 2012) and the Spanish equivalent (LexTALE-ESP; Izura et al., 2014) to ensure high-proficiency in both languages. As previous studies suggest that proficiency may play an important role in L2 semantic processing (e.g., Elston-Güttler, Paulmann, & Kotz, 2005a), the current study only included bilinguals with proficiency scores of above 70% in both LexTALE tests. None of the participants that took part in the LDT took part in the normative surveys.

Materials.

A total of 102 English words were presented as targets for lexical decision, together with 102 non-words pronounceable in both languages. Non-words were matched to the 102 words according to length, and both phonological and orthographic neighbourhood sizes ($p > .05$; see Appendix 9). Neighbourhood size averages were collected from CLEARPOND data base (Marian et al., 2012) for all words and non-words.

Design and Procedure.

The entire experiment was conducted in the participants' L2 (English; see Appendix 10 for materials). Stimulus presentation and the recording of RTs and accuracy were performed using E-prime (Version 2.0; Psychology Software Tools, Pittsburgh, PA). The trials were distributed across two blocks to allow participants to have a break in the middle of the task, there was no time limit on the break, participants resumed the task when they were ready. The trials in each block were randomised through E-prime. At the beginning of each trial, a blank screen appeared for 1000ms followed by a fixation stimulus (+) that appeared in the middle of screen for 1000ms. The fixation disappeared, and a target (word or non-word) was presented in uppercase letters also in the middle of the screen. The target remained on the screen until the participant responded. However, if the participant did not respond within 3000ms the target was removed from the screen and the next trial began immediately. Lexical decisions were made by pressing the *x* and the *m* keys on a keyboard. Response hand was controlled such that all participants made the *word* response with their dominant hand. Instructions and targets were presented in black, size 18 Courier New font, on a silver background. After the task was finished, the participants completed a language background questionnaire, proficiency tests, and were asked to identify any experimental words they were unfamiliar with. The study lasted approximately 30 minutes.

Results

To identify how bilinguals performed in the LDT, RTs and accuracy were analysed separately. Before the analyses of the RTs could be conducted, the data had to be cleaned. For both words and non-words only correct responses were used, therefore, any incorrect responses (12.78% of the data; non-words accounted for 7.27% of these errors) were removed. Furthermore, any RTs that were below 200 ms and exceeding a criterion of +/-2.5 SD for an individual participant's mean (3.40% of the data) were excluded from the analysis. In addition, any homograph meanings that were unknown to the participant were also excluded from the analysis (0.17% of the data).

Three types of analyses were carried out on the data and are reported below: first, the descriptive statistics seen in Table 3, secondly, a comparison between identical and near-identical interlingual homographs in the LDT; then lastly, a regression aiming to establish what linguistic properties significantly accounted for the bilinguals' results.

Lexical Decision Task.

Descriptive statistics for the RTs and error split between identical and near-identical interlingual homographs can be seen in Table 3. For the purposes of all analyses reported in Study 1b, RTs were log transformed to account for the positive skew seen in the histograms and the skewness statistic being above ±1 (see Appendix 11). Table 3 demonstrates that bilinguals appear faster in the near-identical condition in comparison to the identical condition; and make more errors in the identical condition, compared to the near-identical condition. However, repeated measures t-tests show that there were no significant differences between identical and near-identical interlingual homograph trials in lexical decision RTs (p >.05); but there was a significant difference in accuracy (t (3058) = -5.93, p < .001) with more errors in the identical compared to near-identical (see Appendix 12).

Table 3.
Mean RTs (SD) and percentage of errors (SD) for identical and near-identical interlingual homographs, and for the whole set together.

	RTs (*SD*)	Error Percentage (*SD*)
Identical	806.84 (290.90)	14% (.35)
Near - identical	793.66 (279.74)	8% (.27)
Homograph Combined	799.86 (285.06)	11% (.31)

Note: Log RT not used to allow for an easy comparison and interpretation for the reader.

Correlations: RTs, errors, and linguistic properties.

The aim of Study 1b is to determine the influence of linguistic properties on written word recognition, specifically lexical decision making. Therefore, linguistic characteristics were used as predictor variables to see which ones could predict RTs and accuracy scores. Correlations between means of RTs, percentage of errors, and linguistic properties considered in this study are presented in Table 4.

Table 4.

Correlations between predictor variables, RTs and errors.

	Log RTs	Error Percentage
AoA (English)	**.197****	**-.151****
AoA (Spanish)	**.098****	**.055***
Imageability (English)	NS	NS
Imageability (Spanish)	**-.112****	NS
Bigram Sum (English)	**.184****	**.067****
Bigram Sum with English spelling (Spanish)	**.157****	**.066****
Bigram Sum with Spanish spelling (Spanish)	**.167****	**.078****
Length (English)	**.218****	**.088****
Length (Spanish)	**.217****	**.096****
ON (English)	**-.157****	**-.064****
ON (Spanish)	**-.196****	NS
PN (English)	**-.175****	NS
PN (Spanish)	**-.185****	**-.048***
Frequency (English)	**-.275****	**.165****
Frequency (Spanish)	**-.013****	**-.134****

Note: Log = logarithm. NS indicates that the correlations was not significant; significant correlations are indicated in bold, * *p*<.05; ** *p*<.001.

There were significant *positive correlations* between RTs and AoA ratings, length, bigram frequency measures in both languages, Levenshtein Distance. Interlingual homographs that were learned later in life, were longer in length, and higher in bigram frequency showed longer RTs. These patterns are consistent with the findings of previous research (e.g., Assink et al., 2003; Hirsh et al., 2003; Izura & Ellis, 2004; New et al., 2006). There were significant *negative correlations* between RT and frequency, orthographic and phonological neighbourhood size in both languages, as well as imageability in Spanish. Words with higher frequency scores in English were recognised faster than those with lower frequency scores, which is consistent previous literature (e.g., Connine et al., 1990; Frederiksen & Kroll, 1976; Gerhand & Barry, 1999; Hino & Lupker, 2000; Hudson & Bergman, 1985; Perea & Rosa, 2002; Richardson, 1976; Rubenstein, Garfield, & Millikan, 1970). Words with larger orthographic or phonological neighbourhoods also elicited faster responses compared to smaller orthographic or phonological neighbourhoods. This is consistent with previous research that has found that richer neighbourhoods, whether that be phonological or orthographic, are helpful in lexical decision making (e.g., Andrews, 1989; 1992; Forster & Shen, 1996; Sears et al., 1995; Yates et al., 2004).

For the error data, a greater proportion of correct responses were elicited by interlingual homographs that were longer, higher in bigram frequency in both languages, high in frequency of occurrence in English and learned later in Spanish. These patterns are consistent with previous findings concerning length, and frequency (e.g., Balota et al., 2004; Perea & Rosa, 2002), though bigram patterns are inconsistent with previous studies which found no significant effects (e.g., Westbury & Buchanan, 2002). Participants were also more likely to be accurate in recognising interlingual homographs which had small neighbourhoods and infrequent usage in the non-target L1 Spanish language or were acquired early in English. This pattern is again consistent with previous research (e.g., Perea & Rosa, 2002; van Heuven et al., 1998; Yates et al., 2004).

Reaction Time Analyses.

The technique used to analyse the RT data is known as multilevel or hierarchical model (Miles & Shevlin, 2001). Hierarchical models are linear regressions where variation of groups can be modelled at different levels (Gelman & Hill, 2007). In the present study, the data were structured hierarchically with three levels: one corresponding to the participants, the second to the orthographic nature of the homograph (identical; near-identical), and the third to the predictor variables (i.e., the linguistic properties seen in Table 4 without the Levenshtein Distance). RT was the criterion variable. Each participant was dummy coded to control for participant individual differences. In contrast to a simple regression, hierarchical regressions, allow for the investigation of the predictive power of the predictor variables while accounting for the systematic unexplained variation between the participants. Similarly, in a hierarchical regression, a higher number of predictors can be entered into the multilevel regression analysis without risking the possibility of producing unreliable estimates. Therefore, this regression was run to enable the investigation of all the linguistic variables outlined previously, while also accounting for orthographic overlap and participant RT variation, hence this analysis allows us to see how much predictive power the linguistic variables and orthographic overlap can have on lexical decision RTs.

Collinearity statistics (i.e., tolerance below 0.1 and the Variance Inflation Factor (VIF) above 10; Dancey & Reidy, 2007) were checked for the assumption of multicollinearity, it was found that this was violated for the following four variables:

length (English; Spanish), and bigram (English; Spanish). Therefore, four regression analyses were performed to counteract the violation in multicollinearity. Lengths of the words and bigram scores were entered into the analyses either in terms of English or Spanish spelling, but *not* at the same time. Consequently, analysis 1 contained bigram Spanish and Spanish length; analysis 2 contained bigram English and English length; analysis 3 contained bigram Spanish and English length; and analysis 4 contained bigram English and Spanish length. Regression coefficients are presented below in Table 5 (see Appendix 13 – 16).

Table 5.

Standardised regression coefficients (β), R value, and adjusted R2 for the four multilevel analyses carried out on RTs.

	Analysis 1	Analysis 2	Analysis 3	Analysis 4
Step 2				
Homograph type	-.025	-.025	-.025	-.025
R	.442	.442	.442	.442
R^2	.196	.196	.196	.196
Adjusted R Square	.186	.186	.186	.186
Intercept	2.748**	2.748**	2.748**	2.748**
Step 3				
AoA (English)	.040	.039	.040	.040
AoA (Spanish)	.037	.036	.036	.037
Imageability (English)	.028	.029	.029	.029
Imageability (Spanish)	-.003	-.003	-.002	-.004

Bigram Sum (English)	.041	.032	.038	.035
Bigram Sum (Spanish)				
Bigram Sum with English spelling	-	.015	-	.017
Bigram Sum with Spanish spelling	.008	-	.006	-
Length (English)	-	-.008	-.005	-
Length (Spanish)	-.015	-	-	-.018
ON (English)	.025	.026	.027	.024
ON (Spanish)	**-.203****	**-.205****	**-.203****	**-.205****
PN (English)	-.018	-.016	-.017	-.017
PN (Spanish)	**.168***	**.171***	**.170***	**.169***
Frequency (English)	**-.180****	**-.182****	**-.180****	**-.182****
Frequency (Spanish)	.016	.015	.016	.015
R	.506^{+}	.506^{+}	.506^{+}	.506^{+}
R^2	.256	.256	.256	.256
Adjusted R Square	.244	.244	.244	.244
Intercept	2.812**	2.812**	2.810**	2.815**

Note: Significant variables are reported in bold. * *p* < .05, ** *p* < .01, $^+$ F Change *p* <.01. Step 2: homograph type; step 3: linguistic variables. The bigram sums of English and Spanish spelling is obtained from the Spanish language for the near-identical interlingual homographs, whereas the “bigram sum English” derives from the English language for all interlingual homographs. Intercept derives from the unstandardized coefficient.

As can be seen from Table 5, the adjusted R square, and thus the variance does not change across the four analyses conducted. In analysis 1, the hierarchical regression revealed that at Step 1, there was a significant *F* change (p <.001) and the participants contributed significantly to the model, *R square change* = .195; *F* (29, 2661) = 21.97, *p* <.001, and accounted for approximately 19.5% of the variation in RT. In step 2, the two types of interlingual homographs were introduced, and while the overall model was significant as this included the participants from step 1, *R square change* = .001; *F* (30, 2661) = 21.31, *p* <.001. There was no significant F change (*p* = .156) which would suggest that orthographic overlap, that is whether the interlingual homograph is identical or near-identical, did not significantly impact RTs as there were no significant difference in the variance accounted for between the stimuli. In the last step, introducing the predictor variables to the model explained more with approximately 25.6% of the variation in RT, and contributed significantly to the model, *R square change* = .061; *F* (42, 2661) = 20.97, *p* <.001; and there was a significant F change (*p* <.001).

The strongest predictor was Spanish ON, recording the highest beta values across all four analyses, with the pattern suggesting that RTs were faster as Spanish ONs increased in size. Words with richer orthographic neighbourhoods have previously been found to facilitate the word recognition process in L1 even when the task is conducted in the bilinguals' L2 (e.g., Mulder et al., 2018). Although this finding would suggest that despite the task being in the bilinguals' L2 English, the bilinguals' L1 Spanish was facilitating the recognition of the word. English frequency was also a significant predictor, with the pattern suggesting that RTs were faster as frequency scores increased. Frequency effects are consistent with previous findings in the literature using LDTs (e.g., Hino & Lupker, 2000; Perea & Rosa, 2002). Finally, Spanish PNs also contributed significantly to bilinguals' RTs, with the pattern suggesting that RTs were faster as PN decreased in size. Taken together, the findings from the regression analysis on the LDT, tentatively suggest that bilinguals activate both languages, as evidenced by the significant linguistic characteristics. This finding could therefore provide evidence, that would be in line with the BIA+ model (Dijkstra & van Heuven, 2002) which suggest that all possible word candidates are activated in a lexical task conducted explicitly in the bilinguals' L2.

The predictor variables that were not significant were AoA, imageability, and length in both English and Spanish, bigram scores in all spelling variations, orthographic and phonological neighbourhoods in English, and frequency in Spanish which suggest that these variables did not contribute to the bilinguals RT performance.

The strongest predictor was Spanish orthographic neighbourhood. This finding indicated that, despite the task being in the participants' L2, the native language was facilitating the recognition of the word. Spanish phonological neighbourhood size also contributed significantly to RTs, with the pattern suggesting that RTs were faster as the number of phonological neighbours decreased. English frequency was also a significant predictor, with the pattern suggesting that RTs were faster for higher frequency targets. Frequency effects such as this are consistent with previous findings in the literature using lexical decision tasks (e.g., Hino & Lupker, 2000; Perea & Rosa, 2002). This will be explored further in the discussion. Taken at face value, though, the findings indicate that there is some influence of the native language on word recognition in L2.

Error Analyses.

Logistic multilevel hierarchical regressions were performed to assess the impact of the identified linguistic properties on the likelihood that respondents would correctly identify an interlingual homograph as a word. The number of errors was the criterion variable. Each participant was dummy coded to control for participant individual differences. The multilevel technique allowed the accuracy of each participant for each homograph to be considered, and thus, accuracy was registered as a dummy variable (correct responses were coded as 1, incorrect responses as 0). The analysis was structured hierarchically with a three-level hierarchy: one corresponding to the participants, the second to the orthographic overlap of the interlingual homograph (identical, near-identical), and the third to the predictor variables (i.e., linguistic properties seen in Table 4 except the Levenshtein Distance).

Due to the violation in multicollinearity, four regression analyses were performed. Lengths of the words and bigram scores were entered into the analyses either in terms of English or Spanish spelling, but *not* both at the same time. Consequently, analysis 1 contained bigram Spanish and Spanish length; analysis 2 contained bigram English and English length; analysis 3 contained bigram Spanish and English length; and analysis 4 contained bigram English and Spanish length. A summary of the results from the four analyses can be seen in Table 6 (Appendix 17 – 20).

Table 6.

Wald Statistic and R square for the four multilevel analyses carried out on errors.

	Analysis 1	Analysis 2	Analysis 3	Analysis 4
Step 2				
Homograph type	37.974**	37.974**	37.974**	37.97**
Chi-Square	304.85**	304.85**	304.85**	304.48**
Cox & Snell R^2	.095	.095	.095	.095
Nagelkerke R^2	.191	.191	.191	.191
Intercept	**1.669****	**1.669****	**1.669****	**1.669****
Step 3				
AoA (English)	3.595	3.254	3.484	3.451
AoA (Spanish)	**12.987****	**12.517****	**11.449***	**14.003****
Imageability (English)	3.184	3.946*	3.795	3.297
Imageability (Spanish)	**5.741***	**7.771***	**5.919***	**7.263***

Bigram Sum (English)	2.193	.328	1.690	.781
Bigram Sum with English spelling	-	.611	-	.678
Bigram Sum with Spanish spelling	2.645	-	3.099	-
Length (English)	-	**6.676***	**6.355***	-
Length (Spanish)	**9.830***	**-**	**-**	**10.636***
ON (English)	3.614	4.065*	3.796	3.967
ON (Spanish)	**30.368****	**31.749****	**29.332****	**32.385****
PN (English)	**5.217***	**4.541***	**4.687***	**5.089***
PN (Spanish)	**34.508****	**42.269****	**37.884****	**373.831****
Frequency (English)	**56.259****	**55.893****	**55.661****	**56.336****
Frequency (Spanish)	**55.319****	**53.961****	**56.493****	**53.161****
Chi-Square	612.04**	605.84**	608.404**	305.167**
Cox & Snell R^2	.181	.180	.180	.181
Nagelkerke R^2	.364	.361	.362	.363
Intercept	3.292*	3.939*	3.733*	3.440*

Note: Significant linguistic variables are reported in bold. * $p < .05$, ** $p < .001$. Intercept derives from the unstandardized coefficient.

In analysis 1, the full model containing all predictors was statistically significant, $\chi 2$ (43, N = 3060) = 612.04, $p < .001$, indicating that the model was able to distinguish between bilinguals who made a correct or incorrect lexical decision to the interlingual homograph. The model explained 36.4% of the variance in accuracy (Nagelkerke $R^2 = .364$); Nagelkerke, 1991). In step 1 (i.e., participants only) the model was able to correctly classify 89.1% of the cases, in step 2 where orthographic overlap was introduced (i.e., identical, near-identical) the model was able to correctly classify 89.8% of the cases. Lastly, in step when the linguistic properties were introduced, the overall model was able to correctly classify 90.4% of the cases.

In order to determine the direction of the relationship between the predictor variables and accuracy, odds ratios (ORs) were used. According to Tabachnick, Fidell, and Ullman (2007) the odds ratio represents "the change in odds of being in one of the categories of outcome when the value of a predictor increases by one unit" (p. 461). Therefore, to see which predictor variables made the strongest contribution to the model, the Exp (B) values were looked as these are the ORs (i.e., if below 1 = fewer errors made; if above 1 = more errors made). The strongest predictor was English frequency across the four analyses.

Table 6 shows that the linguistic variables that consistently made a significant contribution to the model across the four analyses were: homograph type, AoA Spanish, imageability in Spanish, ON in Spanish, PN in English and Spanish, and frequency in English and Spanish. All made a significant contribution to the model predicting lexical decision accuracy; however, the first thing to note is that, frequency in both languages had the largest coefficient compared to the other significant predictors, suggesting that this variable made the largest contribution to the bilinguals accuracy performance. Secondly L1 characteristics influenced the bilingual's performance again in an L2 task.

Bilinguals were *less likely* to make errors to interlingual homographs with higher AoA scores in Spanish than those with lower scores. *Fewer errors* were made to those homographs with higher imageability scores in Spanish compared to those with lower scores. *Fewer errors* were made to those homographs with more PNs in Spanish compared to those with lower scores, and *fewer errors* to those homographs with higher frequency scores in Spanish. Thus, it seems that similarly to RTs, the fact that L1 Spanish linguistic variables are significantly correlated with error responses would support the idea that all possible word candidates are activated at the same time regardless of language (Dijkstra & van Heuven 2002).

The finding that bilinguals were *more* likely to make errors to interlingual homographs with higher ON in Spanish, suggest that it may have been confusing for bilinguals to have these neighbours activated when the task was in their L2 English, and this interfered with the lexical decision process. Bilinguals were also *more* likely to make errors to interlingual homographs with higher scores of PNs in English. Unexpectedly, bilinguals were *more likely* to make errors to interlingual homographs which had higher frequency scores in English; one might expect that errors would be less likely to be made on the higher frequent words as they are more familiar.

Length in English (analysis 1 and 3) and Spanish (analysis 2 and 4) made a significant contribution to error responses when they were included in the analysis. Bilinguals were *more likely* to make errors to interlingual homographs which were longer in length compared to those which were shorter.

Finally, AoA in English, English imageability, bigram sum in English, bigram sum with English spelling (analysis 2 and 4), and bigram sum with Spanish spelling (analysis 1 and 3) did not reach significance in any of the analyses suggesting that these variables did not have a significant impact on bilinguals' ability to make a lexical decision. Consequently, as these variables are not consistently making significant contribution to the model, they will not be considered as making a significant contribution to bilinguals’ performance overall.

To conclude, homograph type, AoA and orthographic neighbourhood in Spanish, imageability, frequency and phonological neighbourhood in both languages all made a significant contribution to the model predicting lexical decision accuracy. The first thing to note here is that again it can be observed that there influences of L1 characteristics on performance in an L2 LDT. What is particularly interesting is that there are opposing effects of the variables in English and in Spanish. For example, as the number of *Spanish* phonological neighbours increased our participants were more likely to make errors whilst accuracy was likely to be better as the number of *English* phonological neighbours increased. The same pattern was shown in relation to orthographic neighbours and to frequency. Finally, AoA in English, bigram sum in English, bigram sum with English spelling, and bigram sum with Spanish spelling did not reach significance in any of the analyses suggesting that these variables did not have a significant impact on bilinguals’ ability to make a lexical decision response.

However, in this stimulus list there were near-identical interlingual homographs with different Levenshtein Distance spellings, but spelling was not explicitly explored in this analysis. In the cognate literature, facilitations effects are moderated by the level of orthographic overlap, as spelling differences increase, the facilitations effects decrease. However, orthographic overlap has not been investigated in interlingual homographs, therefore an open question remains whether spelling variations can modulate the L1 and L2 effects in interlingual homograph processing. Therefore, the next section aims to explore the effect of variations in orthographic overlap through a new set of analyses focusing on near-identical interlingual homographs only, to see whether this variations in spelling affects bilinguals' performance.

Near-Identical Reaction Time Analysis.

Homographs spelled identically in both languages will not have a Levenshtein Distance value because they are identical, therefore this analysis only investigates near-identical interlingual homographs and their spelling variations. To recap, most of the near-identical interlingual homographs had a score of 1 (*n* = 34), followed by 2 (*n* = 9), 3 (*n* = 6), 4 (*n* = 1), and the maximum Levenshtein Distance being 5 (*n* = 2). Therefore, in order to investigate whether spelling variation in near-identical interlingual homographs affect lexical decision processing, an additional analysis was performed. The data were structured hierarchically with two levels: one corresponding to the participants, and the second to the predictor variables (i.e., the linguistic properties seen in Table 4).

Collinearity statistics (i.e., tolerance below 0.1 and VIF above 10; Dancey & Reidy, 2007) were checked for the assumption of multicollinearity, it was found that this was violated for the following six variables: length (English; Spanish), bigram (English spelling; Spanish spelling), and Spanish neighbourhood (ON; PN). Consequently, eight regression analyses were performed to counteract the violation in multicollinearity. Lengths of the words and bigram scores were entered into the analyses either in terms of English or Spanish spelling, but *not* at the same time; and Spanish neighbourhood, either ON or PN, were entered into the analysis. Therefore, analysis 1 contained English length, bigram English, and ON; analysis 2 contained English length, bigram English, and PN; analysis 3 contained English length, bigram Spanish, and ON; analysis 4 contained English length, bigram Spanish, and PN;

analysis 5 contained Spanish length, bigram English and ON; analysis 6 contained Spanish length, bigram English and PN; analysis 7 contained Spanish length, bigram Spanish, and ON; and finally, analysis 8 contained Spanish length, bigram Spanish, and PN. Regression coefficients are presented below in Table 7 (see Appendix 21 – 28).

Table 7.

Standardised regression coefficients (β), R value, and adjusted R2 for the four multilevel analyses carried out on RTs for near-identical interlingual homographs only.

	Analysis 1	Analysis 2	Analysis 3	Analysis 4	Analysis 5	Analysis 6	Analysis 7	Analysis 8
Step 2								
AoA (English)	.044	.044	.055	.054	.044	.045	.055	.054
AoA (Spanish)	**.106***	**.112***	**.105***	**.110***	**.108***	**.112***	**.106***	**.109***
Imageability (English)	.026	.022	.020	.017	.025	.022	.020	.018
Imageability (Spanish)	**.101***	**.100***	**.102***	**.101***	**.101***	**.101***	**.103**	**.101***
Bigram Sum (English)	-.034	-.029	.019	.019	-.034	-.028	.022	.022
Bigram Sum with English spelling	.059	.047	-	-	.064	.049	-	-
Bigram Sum with Spanish spelling	-	-	-.016	-.020	-	-	-.016	-.022
Length (English)	-.009	.001	.013	.019	-	-	-	-
Length (Spanish)	-	-	-	-	-.022	-.005	.008	.021

ON (English)	.004	-.003	.015	.007	.002	-.004	.014	.007
ON (Spanish)	-.045	-	-.037	-	-.048	-	-.037	-
PN (English)	.036	.029	.020	.015	.035	.028	.109	.015
PN (Spanish)	-	-.015	-	-.011	-	-.016	-	-.009
Frequency (English)	**-.227****	**-.233****	**-.216****	**-.223****	**-.227****	**-.233****	**-.2016****	**.006****
Frequency (Spanish)	-.003	.004	.003	.008	-.003	.004	.003	.006
Levenshtein Distance	**.110***	**.110***	**.104***	**.104***	**.113***	**.111***	**.105***	**.104***
R	.559^{+}	.558^{+}	.559^{+}	.558^{+}	.559^{+}	.558^{+}	.559^{+}	.558^{+}
R^2	.312	.312	.312	.312	.312	.312	.312	.312
Adjusted R Square	.291	.291	.291	.290	.291	.291	.291	.290
Intercept	2.874**	2.867**	2.856**	2.853**	2.878**	2.852**	2.856**	2.878**

Note: Significant linguistic variables are shown in bold. * $p < .05$, ** $p < .01$, $^{+}$ F Change $p < .01$. Step 2: linguistic variables including Levenshtein Distance. The bigram sums of English and Spanish spellings are obtained from the Spanish language, whereas the bigram sum English were derived from the English language. Intercept derives from the unstandardized coefficient.

As can be seen from Table 7, the adjusted R square, and thus the variance does not drastically change across the eight analyses conducted. In analysis 1, the hierarchical regression revealed that at Step 1, there was a significant F change (*p* <.001) and the participants contributed significantly to the model, *R square change* = .220; *F* (29, 1409) = 13.40, *p* <.001, and accounted for approximately 20.3% of the variation in RT. Introducing the predictor variables to the model explained more with approximately 29.1% of the variation in RT, and contributed significantly to the model, *R square change* = .093; *F* (42, 1409) = 14.79, *p* <.001; and there was a significant F change (*p* <.001).

The strongest predictor was English frequency, it has the highest beta value across all eight analyses, with the pattern showing that RTs were faster as frequency scores increased, which is consistent with previous findings in the literature (e.g., Hino & Lupker, 2000; Perea & Rosa, 2002). Spanish AoA also effected RTs, with the pattern suggesting that bilinguals took longer to respond to homographs with higher AoA scores. Moreover, Spanish imageability also influenced RTs, with RTs taking longer for words that had higher imageability scores, this may be because higher imageability scores are easier to access and therefore the L1 interferes more. Consequently, because the task is the bilinguals' L2 English, this imageability effect is reversed and interferes with the task. Finally, the Levenshtein distance was also significant, with those near-identical homographs with higher values (i.e., more spelling differences) producing longer RTs.

The predictor variables that were not significant were AoA and imageability in English, length, ON and PNs in both languages, Spanish frequency, and the three bigram variations.

Near-Identical Error Analysis.

A logistic multilevel hierarchical regression was performed to assess the spelling of near-identical interlingual homographs and various linguistic properties, and their impact on the likelihood that respondents would correctly identify a near-identical interlingual homograph as a word. The analysis was structured hierarchically with a two-level hierarchy: one corresponding to the participants, the second to the predictor variables.

Due to the violation in collinearity statistics, eight logistic regression analyses were performed. Lengths of the words and bigram scores were entered into the analyses either in terms of English or Spanish spelling, but *not* at the same time; and Spanish neighbourhood, either ON or PN, were entered into the analysis. Therefore, analysis 1 contained English length, bigram English, and ON; analysis 2 contained English length, bigram English, and PN; analysis 3 contained English length, bigram Spanish, and ON; analysis 4 contained English length, bigram Spanish, and PN; analysis 5 contained Spanish length, bigram English and ON; analysis 6 contained Spanish length, bigram English and PN; analysis 7 contained Spanish length, bigram Spanish, and ON; and finally, analysis 8 contained Spanish length, bigram Spanish, and PN. A summary of the results from the eight analyses can be seen in Table 8 (see Appendix 29 – 36).

Table 8.

Wald Statistic and R square for the eight multilevel analyses carried out on errors.

	Analysis 1	Analysis 2	Analysis 3	Analysis 4	Analysis 5	Analysis 6	Analysis 7	Analysis 8
Step 2								
AoA (English)	**17.206****	**16.832****	**17.916****	**17.342****	**18.196****	**17.482****	**18.668****	**17.773****
AoA (Spanish)	**15.140****	**17.179****	**13.145****	**15.147****	**15.968****	**17.210****	**13.576****	**14.805****
Imageability (English)	1.215	.716	.435	.179	1.084	.660	.381	.164
Imageability (Spanish)	**17.639****	**17.179****	**11.186***	**11.278***	**18.475****	**18.620****	**11.857***	**11.669***
Bigram Sum (English)	.365	.277	1.451	1.042	.589	.391	1.743	1.148
Bigram Sum with English spelling	3.189	**4.359***	-	-	3.147	**4.389***	-	-
Bigram Sum with Spanish spelling	-	-	**8.898***	**11.278***	**-**	**-**	**8.366***	**9.635***
Length (English)	.745	.242	.363	.071	-	-	-	-
Length (Spanish)	-	-	-	-	2.087	.691	1.126	.189
ON (English)	.709	1.019	1.223	1.540	.698	.996	1.220	1.533

ON (Spanish)	2.784	-	2.456	-	1.143	-	1.141	-
PN (English)	.291	.486	.647	.875	.440	.604	.801	.945
PN (Spanish)	-	**7.110***	-	**6.751***	-	**4.525***	-	**4.732***
Frequency (English)	**13.430****	**15.453****	**14.778****	**17.041****	**12.253****	**14.579****	**13.796****	**16.386****
Frequency (Spanish)	**21.895****	**21.943****	**23.455****	**23.176****	**20.183****	**20.949****	**21.938****	**22.408****
Levenshtein Distance	3.684	**4.141***	**6.151***	**6.984***	**4.756***	**4.667***	**6.876***	**7.121***
Chi-Square	272.12**	276.51**	278.53**	282.86**	273.49**	276.97**	279.30**	282.98**
Cox & Snell R^2	.160	.162	.164	.166	.161	.163	.164	.166
Nagelkerke R^2	.384	.390	.392	.398	.386	.390	.393	.398
Intercept	14.085**	14.353**	12.651**	12.761**	12.609**	12.374**	13.665**	14.042**

Note: Significant linguistic variables are shown in bold. * $p < .05$, ** $p < .001$. Intercept derives from the unstandardized coefficient.

In analysis 1, the full model containing all predictors was statistically significant, $\chi 2$ (42, $N = 1560$) = 272.12, $p < .001$, indicating that the model was able to distinguish between bilinguals who made a correct or incorrect lexical decision in the near-identical interlingual homographs. The model explained 38.4% of the variance in accuracy (Nagelkerke $R^2 = .384$; Nagelkerke, 1991). In step 1 (i.e., participants only) the model was able to correctly classify 93% of the cases, in step 2 where the linguistic properties were introduced, the overall model was able to correctly classify 93.7% of the cases.

Table 8 shows that the strongest predictor across the eight analyses was Spanish frequency with a large coefficient. English frequency, Spanish imageability, and AoA in English and Spanish also significantly contributed to all eight models. Bilinguals were *less likely* to make errors to near-identical interlingual homographs with higher frequency scores, to those with higher imageability scores in Spanish, and higher Spanish AoA scores. This effect was consistent in the English AoA scores, as bilinguals were also *less likely* to make errors to those words with higher English AoA scores. This contrasts to what was expected, which was that those with lower AoA scores would have fewer errors as they are more familiar to the participant (e.g., Brysbaert et al., 2000).

Unexpectedly, for English frequency, bilinguals were *more likely* to make errors to those words with higher frequency scores, because the task is in the bilinguals L2, you would expect that more frequent words would lead to fewer errors. Bigram sum with Spanish spelling was significant across all four analyses it was included in, so bilinguals were *more likely* to make errors to interlingual homographs which had a higher bigram sum frequency within the Spanish language. Bigram sum with English spelling was significant in analysis 2 and 6, but not in 1 and 5, therefore this effect is not consistent and will not be considered as contributing to a bilinguals' accuracy.

Finally, English imageability, bigram English sum, Levenshtein Distance (not significant in analysis 1), and length, and neighbourhoods (ONs; PNs) in English and Spanish did not have a significant impact on the bilinguals' accuracy performance.

Discussion

The present study had two aims, firstly to collate a list of both identical and near-identical Spanish – English interlingual homographs, and to describe them in terms of key lexical and sub lexical characteristics. Near-identical interlingual homographs were also validated in a survey to obtain an idea of how confusable to English and Spanish lexical form are with one another. Secondly, given that a variety of linguistic properties have been shown to affect word recognition in L1 and in L2, we examined the impact of these characteristics on the processing of interlingual homographs on a LDT. Regression analyses were performed to examine the impact of orthography (that is, whether the stimuli were, identical or near-identical interlingual homographs), and which linguistic factors affected the lexical decision RT and accuracy of participant responses.

To this end, two surveys were conducted to obtain AoA and imageability ratings in both English and Spanish, and collated linguistic properties for a set of identical and near-identical interlingual homographs, before investigating the impact of these properties on L2 visual word recognition. The analyses concerning the impact of linguistic properties on bilingual word recognition revealed that there were L1 Spanish characteristics that were significant predictors of lexical decision performance in a bilinguals' L2 English, particularly in relation to accuracy, and that several of the variables that commonly influence word recognition were not significant here.

The results from the lexical decision study 1b showed significant influences of Spanish orthographic neighbourhood on RT, of Spanish frequency, AoA and imageability on accuracy, and of Spanish phonological neighbourhood on both RT and accuracy even though the task was conducted entirely in English. Also of interest is the finding that the spelling of the near-identical interlingual homographs had an impact on lexical decision accuracy. However, to our knowledge, there is no previous study that has explicitly explored the relationship between the spelling of near-identical interlingual homographs and their effects on lexical decision processing. Subsequently, it was difficult to predict whether variations in orthographic overlap would affect a bilingual's performance. However, based on the cognate literature it would be expected that identical and near-identical interlingual homographs would both demonstrate *the same pattern*, as generally both identical and near-identical cognates show the same facilitation effects compared to controls (e.g., van Assche et al., 2011; Dijkstra et al., 2010; Peeters et al., 2013).

In line with the findings in the cognate research, no differences between identical and near-identical interlingual homographs were found in the lexical decision RTs. However, more errors were made to identical interlingual homographs than to near-identical interlingual homographs. One potential explanation for these findings is that the task required participants to decide whether the targets were existing *English* words, and therefore because the identical interlingual homographs share orthographic form these are words in both languages, whereas, for the near-identical interlingual homographs the Spanish spelling is not a word in English. Thus, if participants are making their lexical decisions based on the orthographic form, the identical interlingual homographs could be confused for Spanish words, whereas this is unlikely to happen in the near-identical interlingual homographs as the spellings are different between languages. Dijkstra et al. (1998) demonstrated that the processing of interlingual homographs was influenced by task context and by task instructions. Their participants were slower to respond to interlingual homographs when the stimulus set also included words from the non-target language, Dutch, but not when all the words were in English. This might suggest some type of post-access confirmation of which language the word belonged to was taking place. In this study, all the targets were English words though of course half of them were also identical to Spanish words. These identical interlingual homographs may have been sufficient to implicitly alter the context for the participants and trigger a check of which language the stimulus belonged to as part of the processing. Under these circumstances, identical interlingual homographs may have been incorrectly rejected as English words because the participant recognised them as Spanish words instead. This would be particularly likely if the characteristics of the stimulus were such that the Spanish word was easier to access in the participant's lexicon than the English word.

In order to examine this possibility further, the differences between Spanish and English frequency estimates were calculated for identical interlingual homographs (Spanish minus English) and correlated this score with accuracy. There was a significant negative correlation ($r = -.24$, $p < .001$) suggesting that words that were higher in frequency in Spanish than in English were more likely to be incorrectly rejected in the lexical decision task (see Appendix 37). Indeed, the regressions predicting accuracy that were conducted indicated significant influences of Spanish AoA, imageability and frequency with lower accuracy being predicted for words that would usually be considered “easier” to recognise (high frequency, high imageability, early acquired). This is consistent with the explanation offered above and indicates that there is some activation of the participants’ L1 even when performing a task in L2. The findings from this study constitute only tentative evidence for this account, however, and research could specifically examine this in future by varying task instructions to allow “yes” responses for stimuli that are words in either language, which ought to result in an accuracy advantage for identical interlingual homographs.

Moreover, this finding that more errors are made to identical interlingual homographs compared to near-identical interlingual homographs could be explained by the BIA+ model (Dijkstra & van Heuven, 2002) through lateral inhibition which suggests that interlingual homographs are represented by two, with the possibility of partial overlap, competing semantic and lexical representations which interfere with each other. While semantics may not have been relied upon when making lexical decisions, there is the potential that the *lexical* representations of the identical interlingual homographs are competing with one another. Therefore, the reason why bilinguals make more errors to identical interlingual homographs compared to near-identical ones is explained by the BIA+ model (Dijkstra & van Heuven, 2002) as a disproportionately strong effect of lateral inhibition. Subsequently, because the representations are identical, this competition is stronger compared to near-identical interlingual homographs where the lateral inhibition effect is reduced. Thus, the bilinguals need to inhibit the irrelevant L1 meaning more strongly compared to the near-identical interlingual homographs.

The study also isolated and investigated orthographic overlap in near-identical interlingual homographs. While no predictions could be made, it was expected that based on the cognate literature that any L1 effects would be reduced as orthographic overlap increased (e.g., van Assche et al., 2011; Dijkstra et al., 2010; Peeters et al., 2013). Thus, those near-identical interlingual homographs with an increased spelling variation would show weaker L1 effects compared to those with less spelling variation. A Levenshtein Distance effect was found in both the RTs and accuracy scores in the bilinguals' performance: as the Levenshtein Distance increased, the RTs became longer, and fewer errors were made. While the L1 Spanish may have been activated as evidenced by the significant L1 linguistic properties, it could be that fewer errors were made to those words with longer Levenshtein distances as bilinguals may identify near-identical interlingual homographs with greater Levenshtein distances as two distinct words, and therefore, two orthographic representations which leads to fewer errors as the Spanish orthographic representation is only partially activated. Parallels can be drawn from the error data reported in this Study 1b (i.e., fewer errors were made to those words with longer Levenshtein distances), to that of the near-identical cognate literature; where it is generally found that as the form overlap distance increases, the facilitation decreases (e.g., van Assche et al., 2011; Peeters et al., 2013), and sometimes the effect is even annulled (Comesaña et al., 2012; Comesaña et al., 2015; Dijkstra et al., 2010).

In relation to the linguistic properties of the interlingual homographs, it was predicted that bilinguals would be faster and would make fewer errors for words with higher frequencies and imageability scores, shorter lengths, lower AoA ratings, and more ONs and PNs in English. For the Spanish linguistic properties there were two possibilities. There would either be an effect of the L1 properties which would suggest that languages are integrated and selected non-selectively (e.g., Dijkstra & van Heuven, 2002) or alternatively, no effects would be observed, which would suggest that bilinguals are able to enforce some level of control over their languages, as proposed by the IC Model (Green, 1998; 2003). The results of Study 1b indicate that for RTs, there were three linguistic properties that made a significant contribution: English frequency, Spanish ON, and Spanish PN; and for accuracy, there were twelve linguistic properties that made a significant contribution: homograph type, AoA Spanish, English ON, imageability, length, frequency, and PN in both English and Spanish. Each of these findings will now be discussed in turn in more detail.

The results indicated that as English frequency increased, lexical decision times and errors decreased, which was as expected and in line with previous findings that have found there are faster RTs and fewer errors in high frequency words in LDTs (e.g., Connine et al., 1990; Frederiksen & Kroll, 1976; Gerhand & Barry, 1999; Hino & Lupker, 2000; Hudson & Bergman, 1985; Perea & Rosa, 2002; Richardson, 1976; Rubenstein et al., 1970). It is proposed that high frequency words are easier to retrieve because they are used more often (e.g., Brysbaert et al., 2011; Frederiksen & Kroll, 1976; Hino & Lupker, 2000; Perea & Rosa, 2002). Interestingly, Spanish frequency did not influence the RTs, which would support the notion that bilinguals are able to enforce some level of control over their languages as proposed by the IC model (Green 1998; 2003). However, Spanish frequency was significant in the accuracy data, suggesting activation of the bilinguals' L1 Spanish and providing support that languages are activated non-selectively in the error data only (Dijkstra & van Heuven, 2002). Words with higher frequency scores in Spanish had fewer errors than words with lower frequency scores, so, if the notion that all possible word candidates are activated regardless of language is true (e.g., Dijkstra & van Heuven, 2002), it would make sense that those with higher frequency scores, are easier to access, and thus would be facilitatory to the lexical decision process similarly to the words with high English frequency scores. Bilinguals in this task are fundamentally saying "yes" to the target when they find the exact match (to their L2) or when they have enough activation in the lexicon as a whole, and thus making a decision that it is likely a word even if the exact representation cannot be found (Grainger & Jacobs, 1996). While, Spanish frequency is not significant in the hierarchical regressions for RTs, the overall correlation's pattern for Spanish frequency (see Table 4) is negative in the RTs suggesting that as frequency increases the lexical decision time decreases; which fits with the idea that bilinguals use the first representation (whether exact or not) that helps them make their lexical decision (Grainger & Jacobs, 1996), in this case the bilingual’s English L2 being the first representation accessed. Overall, English frequency drives the speed of the lexical decision making when the L2 target is on the screen, which is logical as the LDT is in the bilinguals’ L2.

Orthographic and phonological neighbourhoods have been found to affect word recognition in LDTs, both within and between languages (e.g., Andrews 1997; Coltheart et al., 1977; Forster & Shen, 1996; Luce et al., 1990; Sears et al., 1995; van Heuven et al., 1998). The results from this Study indicate that there were no effects of L2 English ONs or PNs on the RTs; though, both L1 Spanish ON and PN were significant contributors for RTs patterns, in that, as ONs increased, lexical decision times decreased; whereas, for Spanish PN, as they increased so did the lexical decision times. Further, in the accuracy patterns, ONs and PNs in both English and Spanish were found to be significant contributors to the error data, in that as the neighbourhood increased, fewer errors were made. Whilst the English ON finding does not follow the pattern of previous research, the L1 Spanish ONs were found to influence the bilingual’s performance, despite the task being in the L2 English. This would suggest that languages are activated simultaneously, and bilinguals in turn do not have complete selective-access over their languages which would support models such as the BIA+ (van Heuven & Dijkstra, 2002). It is also consistent with research outlined in the introduction which finds L1 neighbourhood effects in tasks that are in a bilinguals' L2 (e.g., Mulder et al., 2018; van Heuven et al., 1998). This Study's findings would indicate that bilingual word recognition is influenced by the number of ONs in the non-target language, but the target language ONs does not significantly influence a bilinguals' performance.

Unexpectedly English PNs did not influence RTs, which is inconsistent with previous research which has found that words with many PNs are easier to recognise and produce faster RTs (e.g., Chen et al., 2011; Yates, 2005; Yates et al., 2004). However, English PNs did influence accuracy scores, in that bilinguals made fewer errors to those interlingual homographs with a higher number of English PNs. This accuracy pattern is consistent with previous research (e.g., Chen et al., 2011; Yates, 2005; Yates et al., 2004). As discussed above in relation to frequency, it could be plausible that bilinguals are using the first representation that helps them make a lexical decision (Grainger & Jacobs, 1996), so instead of an orthographic representation, it would be a phonological representation aiding the decision making and enabling fewer errors to be made.

The length of a word, that is how many letters it has, has been found to influence individual's performance in word recognition (e.g., Balota et al., 2004; New et al., 2006). While a positive correlation was found for RTs, indicating that longer words would result in longer RTs; in the hierarchical regressions, no significant effects were found for length in the bilinguals' performance in RTs. However, in terms of accuracy, bilinguals did make more errors to words with longer lengths which is consistent with previous research that has found that errors increase as the length of a word does (e.g., Balota et al., 2004; New et al., 2006).

Bigram frequency has also been reported to affect tasks involving word recognition, particularly to low frequency words (e.g., Briederman, 1966; Broadbent & Gregory, 1968; Conrad et al., 2009; Massaro & Cohen, 1994; Owsowitz, 1943; 1963; Rumelhart & Siple, 1974; Westbury & Buchanan, 2002); though, some researchers have suggested that bigram frequency does not have an effect on LDTs (Andrews, 1992). In the case of the results in this chapter, bigram frequency did not impact bilingual's performance in either the RT or the accuracy data which is consistent with some of the literature (e.g., Andrews, 1992; Keuleers et al., 2012). As we have seen above, bigram effects are generally present in low-frequency words as opposed to high-frequency ones, therefore the non-significant bigram frequency finding reported in this chapter could be explained by the fact that there is a mixture of low and high-frequency words. Therefore, to fully explore this, future research could use variations of the LDT and have one task with low-bigram frequency words, and the other with high-bigram frequency words to see if the non-significant effects are indeed due to bilinguals being presented with words with a mix of bigram frequencies.

RTs and accuracy were not affected by AoA in English, and this is inconsistent with previous studies that have found that lower AoA scores are associated with faster RTs and lower accuracy scores (e.g., Assink et al., 2003; Bonin et al., 2006; Brysbaert & Cortese, 2011; Brysbaert et al., 2000; Catling et al., 2008; Hirsh et al., 2003; Holmes et al., 2006; Juhasz, 2005; Izura & Ellis, 2004; Monaghan & Ellis, 2002; Richards & Ellis, 2009). Spanish AoA also did not affect RTs, though it influenced accuracy scores, as bilinguals made fewer errors to homographs with higher AoA scores. This is surprising given that it would be expected that bilinguals would make fewer errors to those words with lower scores as they would be more familiar to the bilingual. However, the task was undertaken in the bilinguals' L2, therefore it could be possible, that those homographs with higher AoA scores are not as familiar and easily-accessible in the bilingual's lexicon and therefore do not interfere as much as those homographs with lower AoA scores.

RTs and accuracy were also not affected by English imageability (in analyses 2 imageability was significant but this was not consistent across the analyses, therefore is omitted from being a consistent effect) and this finding is inconsistent with previous research that has shown that highly imageability words are recognised more quickly than those with low-imageability in LDTs (e.g., Balota et al., 2004; Cortese & Schock, 2013; Kounios & Holcomb, 1994; Kroll & Merves, 1986; Schwanenflugel et al., 1988). It was also predicted that if languages were accessed non-selectively, interference effects from the L1's Spanish linguistic characteristics would be present. In the case of Spanish imageability, lexical decision times were not affected; however, bilinguals did make more errors to homographs with higher imageability scores in Spanish. The accuracy data are consistent with Kiran and Tuchtenhagen (2005) who looked at English – Spanish bilinguals in a semantic priming task and found an effect of imageability in both languages; however, the RTs in the reported analyses are not. Though this may be because of the differences in tasks, this Study utilised a simple LDT, and semantics are consequently not always necessary (e.g., Playfoot et al., 2018; Poort & Rodd, 2019). Whereas, in contrast, the priming task in Kiran and Tuchtenhagen (2005) will be more likely to have accessed words at the semantic level. This notion of not using semantics could be argued to explain why imageability effects were not present in the RT data, however, in the error data Spanish imageability was significant which would suggest that semantics are involved (i.e., more errors to those words that had higher imageability scores). Imageability is associated with easier access to a person's semantic representations, and this finding in the error data would provide support that semantic information is represented in an integrated manger in the bilinguals' lexicon and that that they have non-selective access over their languages which is consistent with the BIA+ model (Dijkstra & van Heuven, 2002).

Two important points need to be borne in mind when evaluating the results of the present experiments to those in the literature. The first and possibly more important one is that the linguistic variables were not *manipulated and controlled*, but follow the constraints of the type of stimuli we were interested in.

Secondly, the LDT did not have any filler target words (i.e., words semantically unrelated to the interlingual homographs), only non-words and interlingual homographs were used as targets. Not having filler words might have impacted performance and altered the salience of the L1 characteristics of the interlingual homographs as after a few trials' bilinguals may have realised the Spanish language applicability to the task (i.e., noticed that the identical interlingual homographs were words in both languages), despite it being in the bilinguals L2. This in turn, would indeed mean that the Spanish language would then need to be actively inhibited by the participants (Green, 1998; 2003). Furthermore, including filler stimuli would have acted as a control comparison condition to use as a baseline and see how the two types of interlingual homographs compare against them. Hence, future research using these stimuli in an LDT should consider introducing filler targets to the research design.

Conclusion

In sum, while previous research has explored the effects of linguistic properties on language processing, none have explicitly looked at the relationship between interlingual homographs with different spelling variations and their lexical characteristics. The results indicate that when bilinguals read interlingual homographs in isolation in their L2, linguistic characteristics from both languages (target and non-target) are activated non-selectively which would be in line with the BIA+ model (Dijkstra & van Heuven, 2002) which suggest that all possible word candidates are activated. This pattern is consistent in both identical and near-identical interlingual homographs, suggesting that both homograph types could be processed in a similar fashion. Further evidence for this claim could be provided by presenting the same task in an L1 Spanish environment and conducting the task in the bilinguals L1. If languages are in fact activated non-selectively, then similar results should be obtained. That is, we would see interference from the bilinguals’ L2 English.

A by-product of this chapter is that it has created a stimulus set, and norming data which will be a valuable resource for researchers studying language processing in Spanish - English bilinguals. Researchers interested in using these stimuli should note that this is not an exhaustive list of Spanish-English interlingual homographs as language is continuously evolving.

In the next chapter, interlingual homographs will not only be investigated in an isolated word paradigm, but at a sentence paradigm level. A sentence paradigm allows for the fuller contribution of semantics and will provide further evidence of the interplay of L1 and L2 in bilinguals' language processing.

Chapter 4 Interlingual homographs as semantic primes in L2 sentence contexts: does orthographic overlap matter?

Introduction

Chapter 3 focused on introducing a novel database of Spanish-English interlingual homographs and on exploring how L1 and L2 linguistic properties of the stimuli affect bilinguals' performance in a LDT. The results indicated that when bilinguals make lexical decisions to interlingual homographs in isolation in their L2 English, linguistic characteristics from both the target L2 and non-target L1 are significant predictors of their performance, and this has been taken to confirm the hypothesis that both languages are activated non-selectively. This chapter extends the previous exploration of interlingual homographs by incorporating an investigation of these stimuli within a sentence and priming context. Study 2 embeds interlingual homographs at the end of a sentence frame that provides a semantic context biased towards their L2 readings. Furthermore, the novelty of the current chapter is that the experiments reported made use of both identical and near-identical interlingual homographs in a systematic way, in a sentence context, something that had not been reported in the literature before apart from the study by Di Betta, Okurowska, and Morgan (2015).

Generally, studies using interlingual homographs in a variety of priming tasks and languages have found that bilinguals take longer to process these words compared to control words (e.g., De Groot, Delmaar, & Lupker, 2000; Dijkstra, De Bruijn, Schriefers, & Ten Brinke, 2000a; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Libben & Titone, 2009; Pivneva, Mercier, & Titone, 2014; Poort et al., 2016; Von Studnitz & Green, 2002). Longer processing times are argued to reflect the non-selective activation of both languages, and the fact that this subsequently requires the bilingual to inhibit the interfering non-target meaning. Models such as the Bilingual Interactive Activation plus (BIA+) (Dijkstra & van Heuven, 2002) argue that, because the lexical forms of interlingual homographs are identical (or similar) across languages, there is stronger competition for interlingual homographs compared to regular words. Thus, the inhibitory effects would suggest that both meanings of an interlingual homograph are activated non-selectively until the relevant meaning is selected. However, previous research investigating interlingual homographs has mostly used these stimuli as the target for lexical decisions (e.g., Dijkstra et al., 1998; Dijkstra, Grainger, & van Heuven, 1999; Poort et al., 2016).

One of the novelties of the current study is that the lexical decision is not made to the interlingual homograph, but to a semantically related target. When making lexical decisions to the interlingual homographs, the bilingual may take longer to process these words because the lexical representations are identical (or similar) to one another (Dijkstra & van Heuven, 2002). Whereas, by asking participants to make lexical decisions to semantically related targets of the interlingual homographs (which are used as primes) we can reduce any lexical overlap effects by focusing on the semantics of the word. In this study bilinguals are required to activate the interlingual homograph’s meanings as cued by a sentential context, and sentences bias the L2 meaning of the interlingual homograph (see Table 9). Therefore, the task investigates whether bilingual participants just access the L2 meaning (as demanded by the task) or whether L1 meaning is also activated. This will be achieved by manipulating the relationship between the prime (interlingual homograph) and the target word that is either semantically related to the L2 meaning, L1 meaning, or is unrelated. Hence, the paradigm used for the experiments reported in this Chapter will help to provide further evidence that the bilingual is non-selectively activating both *meanings* of the interlingual homographs.

A further novelty of this study is that it makes use of both identical and near-identical interlingual homographs, whilst previous literature has mainly focused on the identical orthographic forms. This is an important addition, because, as we saw in Chapter 3 (validation survey in study 1a), there is tentative evidence to suggest that Spanish - English bilinguals do find near-identical interlingual homographs confusing which provides support that these stimuli are important to investigate. Furthermore, near-identical interlingual homographs are interesting because they allow us to explore the role of orthographic overlap in modulating lexical-semantic activation. Studies on cognates have investigated both identical (e.g., PIANO in English and Spanish) and near-identical ones (e.g., the words MAP-MAPA in English and Spanish) words and have found that the overall magnitude of cognate facilitatory effects is moderated by differences in form overlap (e.g., Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Font, 2001). As previously mentioned in Chapter 3, stronger facilitation effects are produced for identical cognates compared to near-identical cognates, suggesting that such effect reduces in magnitude with decreasing form overlap (e.g., Dijkstra et al., 2010; Font, 2001; Peeters, Dijkstra, & Grainger, 2013; van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011). Cognate facilitation effects are commonly accredited to the fact that an L2 cognate also activates its L1 orthographic counterpart as well as the same semantic representation (see Dijkstra et al., 2010, for more information on the representational structure of cognates), which subsequently increases the recognition speed of cognates compared to non-cognates. This effect is evidence in support of a bilingual lexicon that stores languages in an integrated fashion, and that the two representations (L1 and L2) are activated.

While there has been a substantial amount of research on identical and near-identical cognates, the research on near-identical interlingual homographs is limited. One study that has explored orthographic overlap in interlingual homographs is that by

Di Betta et al. (2015) who specifically manipulated the orthographic overlap of Polish-English interlingual homographs presented in a L2 sentence context to investigate the performance of proficient Polish-English bilinguals in a sentence priming with lexical decision task. The interlingual homographs acted as primes at the end of the sentence and were followed by targets that were either related to the English, Polish or unrelated to the meaning of prime. For instance, a near-identical interlingual homograph used in the study is CARAVAN – with KARAWAN meaning *hearse* in Polish, and an example sentence is as follows "Last night we went camping in our caravan". Lexical decision targets would be related to the English meaning (trailer), to the Polish meaning (cemetery), or were unrelated. Although the sentence was biased to the L2 meaning of the interlingual homographs, bilinguals demonstrated priming effects with both English and Polish semantically related targets, thus suggesting activation of both languages. However, results also showed different priming effects depending on the homograph type. For the Polish-related target, identical interlingual homographs demonstrated negative priming (i.e., longer RTs to the Polish-related target compared to unrelated targets), which was interpreted as evidence of inter-language competition and inhibition, whereas the near-identical interlingual homographs produced a positive priming effect (i.e., faster RTs to the Polish-related target compared to unrelated targets). This suggests that the L1 may be less inhibited in this condition. Therefore, Di Betta et al. (2015) proposed that different control mechanisms might be employed for identical and near-identical interlingual homographs. Nevertheless, their results suggest that the bilinguals activated the Polish L1 meaning of the interlingual homograph, despite the task being biased to the L2 meaning of the interlingual homograph. These effects were interpreted in that both languages were activated non-selectively.

Inhibitory effects like the ones observed by Di Betta et al. (2015) for identical interlingual homographs are more likely to be observed when the experiment includes words from the bilingual's non-target language that require a 'no' response (e.g., De Groot et al., 2000; Dijkstra et al., 1998; Von Studnitz & Green, 2002), as discussed in Chapter 3. As an example, Dijkstra et al.'s (1998) experiments demonstrate that task demands (i.e., language specific or generalised; presence of the other non-target language words requiring a 'no' response) can influence the processing of interlingual homographs. This can also be related to the different pattern of results which Di Betta et al. (2015) found for the identical and near-identical interlingual homographs, the task demands could be a contributing factor for the different priming effects observed.

Furthermore, a later study by Dijkstra et al. (1999) argued that the absence of homograph effects in their previous study (Dijkstra et al., 1998) could be because of the orthographic and phonological overlap of the homographs. For instance, homographs can have the same pronunciation in the two languages (e.g., English word PET meaning *cap* in Dutch) or have different pronunciations across languages (e.g., STAGE meaning *internship* in Dutch). Dijkstra et al. (1999) conducted a LDT in Dutch-English bilinguals' L2 and found that RTs were faster and more accurate to identical homographs with identical pronunciations in both languages compared to control words, and that identical interlingual homographs with identical pronunciations were faster compared to interlingual homographs with different pronunciation. Dijkstra et al. (1999) interpreted this phonological inhibition effect (i.e., phonological overlap inducing slower RTs and more errors) as the consequence of activation of the two distinct phonological representations of the written interlingual homograph. Thus, phonological inhibition occurs because after a given letter string activates all compatible phonological codes independently of language (Brysbaert, Van Dyck, & Van de Poel, 1999; Nas, 1983), the activated nonidentical phonological lexical representations may compete at a lexical level (e.g., through lateral inhibition; Dijkstra et al., 1999) and this results in a delayed identification of the item in the target language. Overall, these findings demonstrate the importance of controlling for phonological similarity across languages and suggest that the null results presented in Dijkstra et al.'s (1998) word may have resulted from both the phonological and orthographic overlap between stimuli. Subsequently, the interlingual homographs used in this thesis will all have distinct phonological forms in the two languages (for instance PIE is pronounced /pai/ in English, but /pyeh/ in Spanish). While controlling for phonological overlap cannot guarantee that the bilinguals will process the interlingual homographs in the same manner, it can push the effect to be as consistent as possible. However, phonology is something that we cannot disregard as having an influence on interlingual homograph processing and so the reader should be mindful of this.

While tasks involving isolated words (e.g., LDT and priming paradigms) have been useful in exploring the effects of interlingual homographs at the lexical and semantic level, sentences allow for a more natural exploration of these stimuli as words are rarely encountered in isolation. However, compared to isolated paradigms, evidence of interference between meanings has been more elusive to find in sentence contexts, with some studies not finding any effects, even in low constraint sentences (that is, sentences that are not biased towards a specific meaning of the interlingual homograph; for a review, see Lauro & Schwartz, 2017; Van Assche, Duyck, & Hartsuiker, 2012). For example, Schwartz and Kroll (2006) used a sentence and naming paradigm with proficient Spanish (L1) - English (L2) bilinguals and manipulated the degree of semantic constraint towards the target language; sentences were either weakly or highly constrained toward the L2 English meaning of the interlingual homographs. An example of a low constraint sentence would be “We felt a bit nervous when we saw the *fin* of the shark in the distance.", ‘fin’ means *end* in Spanish; an example of a high constraint sentence would be "From the beach we could see the shark's *fin* pass through the water.". Bilinguals were told they would see sentences in their L2, one word at a time, and that one word in each sentence would appear in red, and that they had to name this word out loud into the microphone as quickly and accurately as possible. This was to ensure that the bilinguals said the interlingual homograph out loud. Bilinguals did not differ in RTs in either the low or high constraint sentences when processing interlingual homographs compared to control sentences. However, those with lower proficiency in L2 English did make more errors on interlingual homographs compared to control words in both sentence types, with a trend towards more errors in the low-constraint condition. Schwartz and Kroll's (2006) results suggest a limited role of sentence frames in eliciting activation of the interlingual homographs competing meanings, particularly for highly proficient bilinguals. Only those bilinguals with lower proficiency made more errors. Subsequently, the presence of a contextual constraint (low and high) allowed for language-selective processing in the Schwartz and Kroll’s (2006) study. While the BIA+ model (Dijkstra & van Heuven, 2002) suggests that language selection follows activation and competition, this finding indicates that competition may not always occur, or it may be resolved quickly. These findings could also be explained from the reordered access principles, where preceding context can strongly bias one candidate over a form-related semantic competitor (e.g., Duffy, Henderson & Morris, 1989; Rayner & Duffy, 1986). The IC model (Green 1998) would explain these results by suggesting that the non-relevant L1 Spanish was successfully and quickly inhibited, and the proficient bilinguals had more control in this task. In contrast, bilinguals with lower proficiency in the L2 English may have had less success in inhibiting their L1 Spanish which subsequently resulted in the increase in errors, particularly in the sentences where the constraint was low and more inhibition of the L1 was required.

More studies with this type of stimuli have found further interference from interlingual homographs in sentence contexts (Hoversten & Traxler, 2016; Jouravlev & Jared, 2014; Lauro & Schwartz, 2017; Libben & Titone, 2009; Pivneva, Mercier, &

Titone 2014; Titone, Libben, Mercier, Whitford, & Pivneva, 2011). For instance, Libben and Titone (2009) analysed French- English bilinguals' eye movements when reading sentences that included interlingual homographs; they used high and low constraint sentences in a similar manner to Schwartz and Kroll's (2006) study. Initial activation (also known as early activation) processes were measured by means of bilinguals' first fixation, gaze duration, and skipping rate, whilst late activation was measured by total reading time. In the low-constraint condition, early activation measures indicated that participants spent more time on interlingual homographs compared to control words, however, in high-constraint sentences, interlingual homographs were processed significantly more slowly than control words in gaze duration only. Thus, there were more indicators of L1 interference in the low-constraint sentences, but activation of both meanings was present in both cases, which Libben and Titone (2009) argue to be evidence of interference from the non-target language. Overall, these results provide evidence for an early non-selective stage of bilingual lexical activation, consistent with the assumptions of the BIA+ model (Dijkstra & van Heuven, 2002).

In addition to sentence context, the language in which an experiment is conducted (language context) can affect the degree of language selectivity observed across experiments (Van Kesteren, Dijkstra, & De Smedt, 2012; Wu & Thierry, 2010). For instance, Elston-Güttler, Gunter, and Kotz (2005b) argued that bilinguals are capable of selective language processing as long as they are fully adjusted to a single language setting, and, therefore, that language context can modulate how much interference is present from the non-target language. This process of adjusting to a monolingual language context is called 'zooming in' on a language. In their study, before the experiment begun, proficient German-English bilinguals watched a short film in either their L1 German or their L2 English. The rest of the experiment was conducted entirely in the participants’ L2 English. The experiment consisted of bilinguals making lexical decisions to a target word following a sentence that ended with a prime word that was either an interlingual homograph or a matched control word. The target word was the L2 English translation of the L1 German meaning of the interlingual homograph. For instance, for the sentence “The woman gave her friend a pretty” the prime would either be GIFT which is a German – English interlingual homograph meaning *poison* in German, or the control word SHELL, and the target would be POISON. Priming of the L1 German target word was only found in those participants who had watched the film in their L1 prior to the task. This provides evidence of activation of their L1, and thus un-selective access unless participants are zoomed into the target language of the task. The priming effects of the L1 German disappeared by the second half of the experiment. Elston-Güttler et al. (2005b) suggest that this is because bilinguals adjusted to the change in language mode and were therefore able to selectively access words in the target language.

More recently, Poort, Warren, and Rodd (2016) provided further support to the idea that recent experience in one language influences processing of the same word-forms such as interlingual homographs in a different language. Proficient Dutch-English bilinguals read L1 Dutch sentences containing interlingual homographs, which were presented again 16 minutes later in isolation in a LDT in the participants’ L2 English. So, the first task was in L1 Dutch, and the second in the L2 English. Activation of the L1 Dutch was found as participants were slower when making lexical decisions to interlingual homographs compared to controls, particularly when switching languages. Poort et al.’s (2016) results show that language switching can influence lexical processing and have residual effects in bilingual speakers at the level of individual words when making lexical decisions. These findings support the importance of language context when conducting research with interlingual homographs. Therefore, in the present study, bilinguals will carry out the experiment in a L2 context, and the experimenter will be a native English speaker with no foreign accent to ensure that participants are ‘zooming in’ (Elston-Güttler et al., 2005b) to the correct target L2 language. Thus, the present study will bias all the aspects which might influence a bilingual’s processing of interlingual homographs. This will present a strong test of the non-selective access hypothesis, because it will investigate whether there is any evidence of the L1 being activated, when the variables outlined above are attempting to promote the activation of the L2 English only.

The present experiment: Study 2

This chapter consists of two main experiments (study 2a and study 2b) and also reports the results of two norming studies that were carried out to validate the stimuli used. All the studies outlined in this chapter received ethical approval from Sheffield Hallam’s University Research Ethics Committee (AM/KW/D&S-345).

Experiment 2a

This is a conceptual replication of Di Betta et al.'s (2015) research as identical and near-identical interlingual homographs were used as the primes for 500 ms at the end of sentences, and sentences were manipulated to convey the L2 English meaning of the interlingual homographs. The same paradigm was used: participants read a written sentence missing a final word, and then were shown the interlingual homograph which completes the sentence (prime), and finally this was followed by a target to which a lexical decision was made. Targets were manipulated to be either related to the L1 Spanish meaning, L2 English meaning, or unrelated in meanings to either language (for an example of stimuli, see Table 9). However, instead of using the Polish language as the L1, the present study involved Spanish - English bilinguals and therefore Spanish will be the L1. While both the current experiment and that reported by Di Betta et al.

(2015) use languages derived from Indo-European heritage, whose orthography relies on the Latin alphabet, and make use of accents in their orthography, there are some differences between the two languages. For example, in the Polish language, the alphabet includes some additional letters with diacritics such as the letter ‘l’ which often has a stroke in it ‘Ł or ł’ (in modern Polish this is usually pronounced /w/; Mazur, 2011) which are not features found in the Spanish language. By using the same L2 English, and changing only the L1 of the bilinguals, one can extend the research by exploring a further language and ascertain how generalizable the previous findings truly are (e.g., Diener & Biswas-Diener, 2016; Shrout & Rodgers, 2018; see Stroebe, 2016, for evidence of why conceptual replications are important to validate previous findings).

Table 9.
Sample stimuli used in Study 2a and 2b.

Homograph Type	Sentence (PRIME)	Target
Identical	My mum finished baking the (PIE)	pastry[1] toe[2] grim[3]
Near-Identical	To unlock the safe you need to know the correct (CODE)	secret[1] arm[2] sheep[3]

Note. [1]Related to the English meaning of the homograph, [2]related to the Spanish meaning of the homograph, [3]unrelated to either meaning. *CODE - CODO, the latter meaning *elbow* in Spanish.

This experiment has two main aims. Firstly, to investigate the interplay of L1 and L2 activation in Spanish – English bilinguals who are living in the UK, and who are taking part in an experiment conducted entirely in their L2 English, with the sentence context biasing the L2 meaning of the interlingual homograph. Thus, the question of interest is to what extent the participants' L1 Spanish is activated during the L2 task, and whether the L1 activation is modulated by sentence meaning. Secondly, the study aims to explore whether there are processing differences between identical and near-identical interlingual homographs. As mentioned earlier, Di Betta et al (2015) showed inhibition of the L1 meaning with identical interlingual homographs, but positive priming for the L1 with near-identical ones.

Based on previous research, it is predicted that bilinguals will show activation of their L1 in the task, but that there will be differences between the priming effects elicited by the interlingual homographs, with positive priming following a near-identical interlingual homographs prime and negative priming following an identical one (Di Betta et al., 2015). Furthermore, based on the cognate literature where facilitation effects are reduced as orthographic overlap is increased (e.g., van Assche et al., 2011; Dijkstra et al., 2010; Peeters et al., 2013), one would expect that bilinguals processing targets primed by identical interlingual homographs will demonstrate stronger priming effects for the L1 Spanish meaning compared to near-identical ones.

The control group will be NEs with no experience of the Spanish language. The control group is important to the design because if the NEs exhibit differences between the unrelated and Spanish related targets, then any priming effects demonstrated among the bilinguals would be redundant as it may not be related to L1 activation, but rather could be due to intrinsic properties of the stimuli. It is predicted that there will be no significant priming for the Spanish related targets: both the unrelated and Spanish related targets will be responded to more slowly than the English-related targets.

Some researchers have found no interlingual homograph effects in lexical decision paradigms (e.g., De Bot, Cox, Ralston, Schaufeli, & Weltens, 1995; Dijkstra et al., 1998; Gerard & Scarborough, 1989; Lemhöfer & Dijkstra, 2004), thus if the bilingual group in the current study display no significant priming for the Spanish related target, this would extend the results to a sentence paradigm and suggest that bilinguals are able to effectively stop their L1 Spanish from interfering in the L2 task and/or that they do not activate the L1 at all. However, if the L1 Spanish does influence the L2 word recognition, then bilinguals should respond differently to Spanish-related targets compared to the unrelated targets (Spanish target priming). This finding would demonstrate that despite the sentences and the experimental environment being biased to the bilinguals' L2, interference from the non-relevant native language is still present.

Study 2 (a)

Method

Participants

A total of twenty-four Spanish-English bilinguals (16 females and 8 males from the Sheffield area and thirty-six NEs (21 females and 9 males) from Sheffield Hallam University served as volunteers. No participants reported a language disability (e.g., dyslexia) or neurological disorders. The participant's characteristics are reported in Table 10.

All bilinguals were educated to at least A-levels or equivalent level, eleven had graduated with a bachelor's degree, and nine had a master's degree. In addition to English and Spanish, ten bilinguals had experience with another language. All NEs were educated to at least A-levels or equivalent, with two participants with a master's degree. Twenty NEs had no experience with another language and sixteen had experience with one other language. However, the inclusion criteria required NEs to *not* be actively learning an L2 and to have no prior knowledge of the Spanish language.

All bilinguals reported a score of 6.5 or higher in the International English Language Testing System (IELTS), which is a test of English language proficiency, which they provided evidence for to the researcher. The test is designed to assess the language ability of non-native English speakers who want to study or work in a country where English is used as the language of communication. Having an overall score of 6.5 is very good and according to the IELTS scoring band, it puts a bilingual somewhere between a “competent user” and “good user”.

All participants were asked to complete a language background questionnaire which asked about their current language experiences and about any previous languages they had tried to learn. Bilinguals reported being dominant in their native language Spanish, though proficient and daily users of the English language (for details, see Table 10). In addition to subjective measures, participants completed the 'Lexical Test for Advanced Learners of English' (LexTALE; Lemhöfer & Broersma, 2012) and bilinguals also did the Spanish equivalent (LexTALE-ESP; Izura, Cuetos, & Brysbaert, 2014) to ensure high-proficiency in both languages. As previous studies suggest that proficiency may play an important role in L2 semantic processing (e.g., Elston-Güttler, Paulmann, & Kotz, 2005a), the current study only included bilinguals with a proficiency score of above 70% in both LexTALE tests (see Table 10).

Table 10.
Participant language characteristics.

	Bilingual	NEs	t-test (*p*)
Age in years (*SD*)	29.0 (*6.3*)	23.0 (*6.9*)	.001
LexTALE	84.4 (*9.4*)	92.2 (*5.6*)	.001
LexTALE-SPA	89.4 (*11.9*)	NA	NA
Living in L2 speaking countries (months)	28.3 (*12.1*)	All their life	NA
Self-rated proficiency English[1]	3.5 (*.5*)	3.8 (*.3*)	<.001
Self-rated proficiency Spanish	4.0 (*.5*)	NA	NA
Daily use English[2]	4.8 (*.5*)	5	Ns
Daily use Spanish	5	NA	NA

Note. NA = not applicable, NEs = Native English Speakers. Standard deviation is reported in brackets unless stated otherwise; some bilinguals had lived in the UK for longer than others as noted by the SD. [1]Self-ratings ranged from 0 to 4 (0 = not at all, 1 = not well, 2 = I can, but with a lot of difficulty, 3 = well, but with a little difficulty, 4 = very well) and were rated on reading, speaking, and understanding; averages are shown. [2]Self-ratings of daily use ranged from 1 to 5 (1 = rarely, 2 = 1-2 times per month, 3 = once a week, 4 = 2-3 times a week, 5 = every day) and were rated on reading, speaking, and understanding; averages are shown. Some of the participants reported some knowledge and use of L3. Independent t-tests were conducted to compare Bilingual and NEs conditions in the language characteristics and can be seen in the third column.

Stimuli

Experimental Sentence Frames

This experiment makes use of 36 identical and 36 near-identical interlingual homographs, selected from the database of 102 stimuli presented in Chapter 3 (see

Appendix 38 for full list of the stimuli and materials). Interlingual homographs were selected as stimuli for Experiment 2a only if they were able to fit at the end of a sentence. For example, *ALAS* (meaning wings in Spanish) is commonly positioned at the start of the sentence in English, not the end, and subsequently was not included in the study. Additionally, because this task was in the bilinguals' L2, we wanted to ensure that the L2 meaning was familiar to them, thus interlingual homographs were selected only if they had a frequency rating that was high enough that L2 speakers of English were likely to know them. Therefore, all interlingual homographs had a frequency value of above 1 on the Zipf scale in both languages (1-3 is labelled as low on the scale; Zipf, 1936; 1949). The frequency measures were taken from SUBTLEX-UK (van Heuven, Mandera, Keuleers, & Brysbaert, 2014); to standardise the frequency scores, the Zipf scale was used (Zipf, 1936; 1949) where frequency ranges from 1 (1 – 3 low frequency words) to 6 (4 – 6 high frequency words) or 7 (function words, pronouns, and verb forms like "have").

The final set included a total of 72 interlingual homographs that were each placed at the end of a sentence frame developed to bias the L2 English meaning. Therefore, the experimental sentences ended in an interlingual homograph which acted as the prime and was followed by a lexical decision target that was either related to the English meaning of the homograph, to the Spanish meaning, or unrelated to either meaning. Table 11 shows an example sentence from each condition. Experimental targets were matched across conditions for frequency, length (number of letters), ONs, and PNs (Orthographic and phonological neighbourhood sizes were taken from the CLEARPOND database; Marian et al., 2012). A one-way ANOVA comparing the properties of the targets in the three experimental conditions was conducted. The reader should note that in the instances where Mauchly's test of sphericity has been violated, Greenhouse-Geisser is reported to correct for the violation of sphericity (Dancey &

Reidy, 2007). The one-way ANOVA demonstrated that there were no significant differences between the targets in terms of **frequency** (English related target *mean* 4.18 (*SD* = .7), Spanish related target *mean* = 4.02 (*SD* = .7), and unrelated *mean* = 4.1 (*SD* = .7); *F* (2, 142) = .82, *p* = .442), **length** (English related target *mean* = 5.5 (*SD* = 1.5), Spanish related target *mean* = 5.4 (*SD* = 1.5), and unrelated *mean* = 5.5 (*SD* = 1.5); *F* (1.81, 128.69) = .08, *p* = .903 *Greenhouse-Geisser corrected*), **ONs** (English related target *mean* = 3.9 (*SD* = 5.7), Spanish related target *mean* = 3.9 (*SD* = 4.4), and unrelated *mean* = 3.8 (*SD* = 4.7); *F* (2, 142) = .04, *p* = .966), and **PNs** (English related target *mean* = 7.7 (*SD* = 9.8), Spanish related target *mean* = 10.0 (*SD* = 11.0), and unrelated *mean* = 7.9 (*SD* = 9.0); *F* (2, 142) = 1.29, *p* = .278; see Appendix 40). Orthographic and phonological neighbourhood sizes were taken from CLEARPOND (Marian et al., 2012).

To obtain a standardised score for the orthographic difference between near-identical spellings in English and Spanish, the Levenshtein Distance was calculated for non-identical interlingual homographs. For the 36 near-identical interlingual homographs the average of the Levenshtein Distance was 1.3 (*SD* = 0.6). Most of the words had a score of 1 (*n* = 27), followed by few with a score of 2 (*n* = 6), and only three with the maximum Levenshtein Distance of 3.

Filler Sentence Frames

Filler sentences were also developed according to the same criteria used for the experimental ones and added to those to make sure that the experimental manipulation was not obvious to participants. Filler sentences could be followed by a word target (*N* = 27) or by a non-word (*N* = 99). In total each participant made lexical decision responses to 198 stimuli (filler *N* = 126, experimental *N* = 72). The filler word and non-word trials did not contain any words semantically related to any meaning of the homographs. Overall, participants made a total of 99 “yes” responses and 99 “no” responses in the LDT.

Using the experimental linguistic properties as a baseline, the filler primes and targets were manipulated and matched to these averages. The filler primes did not significantly differ compared to the experimental primes in length (*t* (26) = -1.99, *p* = .056), frequency (*t* (26) = -1.64, *p* = .113), and neighbourhood size (*t* (26) = -.62, *p* = .541). The filler targets also did not significantly differ compared to the experimental targets in length (*t* (26) = .77, *p* = .449), frequency (*t* (26) = -1.95, *p* = .062), ON (*t* (26) = .55, *p* = .584), and PN (*t* (26) = -.95, *p* = .353; see Appendix 40).

A total of 99 filler sentences were followed by a non-word; using the experimental linguistic properties as a baseline, the filler primes were controlled for length (*t* (71) = -.90, *p* = .369) and neighbourhood size (*t* (71) = -.09, *p* = .930). In addition, filler non-word targets were also matched to experimental targets and did not differ significantly in length (*t* (98) = .06, *p* = .956) and neighbourhood size (*t* (98) = 1.98, *p* = .051; see Appendix 41). A frequency score could not be obtained for this condition as they were non-words.

Table 11.

Example sentences of the experimental and filler targets.

		Target		
Experimental sentence	**PRIME**	English	Spanish	Unrelated
The first concert I attended was held in an	ARENA[1]	stadium	beach	slim
The sea captain ordered his men to lower the	ANCHOR[2]	ship	width	fame

Filler sentence	**PRIME**	**Target**
My sister has many freckles on her	FACE	cheek[3]
My family wants to buy a plot of land to build a	HOUSE	paper[4]
Every Sunday we have a picnic in the	PARK	thack[5]

Note: [1] = identical interlingual homograph meaning *sand* in Spanish, [2] = near-identical interlingual homograph, ANCHO meaning *wide* in Spanish, [3] = filler sentence related target, [4] = filler sentence with unrelated target, [5] = filler sentence with non-word as a target (there were no trials in which an interlingual homograph ended a sentence and preceded a “no” response, only filler sentences were preceded a non-word).

4.1 Semantic Relatedness and Sentence Comprehension Norm Surveys.

Latent Semantic Analysis (LSA; Landauer & Dumais, 1997) was used to measure the relatedness of the experimental prime-target pairs**.** LSA is a technique in natural language processing that analyses relationships between pairs of words by producing values through an online database (Landauer, Foltz, & Laham, 1998). Word pairs with values close to 1 are considered semantically similar, while pairs with values closer to 0 are semantically dissimilar. LSA was used to ensure that the English meaning of the interlingual homographs were related to the English target word and not related to the Spanish target word. Hence, a target that was related to the English meaning of the prime would be scores closer to 1 in LSA, and a target related to the Spanish meaning of the target would be unrelated and therefore scoring closer to 0. This was the case for almost all the pairs. However, there were seven pairs of words that were unable to be processed by the LSA system because they did not exist as pairs in the database. In order to overcome this limitation and to provide a subjective measure of relatedness for all prime-target pairs, two surveys were carried out, one with native English speakers and the other with native Spanish speakers. Spanish bilinguals were asked to access the Spanish meaning of the interlingual homographs (prime) and the NEs the English meaning.

Eight monolingual NEs were recruited for the English relatedness survey (5 females and 3 males; *mean age* = 27 years, *SD* = 7.6). A total of 3 participants had no experience with another language, 4 had experience with one, and 1 had experience with more than one language. Those participants with knowledge of another language stated that they were no longer actively learning an L2 and had no knowledge of the Spanish language.

Eight native Spanish speakers were recruited for the Spanish relatedness survey (7 females and 1 male; *mean age* = 27 years, *SD* = 9.2). A total of 2 participants had no experience with another language apart from English and, 5 had experience with an additional language to English and Spanish and 1 had experience with more than three languages in addition to English and Spanish. All sixteen participants reported a no known language disability (e.g., dyslexia) or neurological disorders.

The English relatedness survey asked NEs to rate pairs of words as being either related or unrelated with each other. Ideally, NEs should rate the interlingual homographs as related to the English target word and unrelated to the Spanish (as they did not know the Spanish meaning of the interlingual homographs). In contrast, the bilinguals were asked to access the Spanish meaning of the interlingual homographs (inhibit the English meaning). Thus ideally, participants would rate the interlingual homographs as related to the Spanish target word and unrelated to the English target word. For example, for the word RED (meaning *network* in Spanish) NEs participants should rate RED-WIFI as being unrelated and RED-BLUE as being related. Whereas Spanish bilinguals should rate it in the opposite way, hence, RED-WIFI would be a related pair and RED-BLUE would be unrelated. On average, the survey took 15 minutes to complete (see Appendix 42 for materials).

Furthermore, NEs were also asked in the survey to rate the 72 experimental sentences that would be used in the main experiment in terms of comprehension and whether the last word (i.e., interlingual homograph) fitted within the sentence on a scale ranging from 1 (very good) to 5 (very poor). For example, 'The strawberries in the supermarket were a shade of red'. The average rating for sentence comprehension was 1.2 (*SD* = 0.4); 78% of sentences were rated as very good and 22% of sentences were rated as good.

The analysis of the English relatedness surveys' responses showed that there were two pairs of words that were identified as being problematic (i.e., PAN, meaning *bread* in Spanish, paired with FLOUR; LAME meaning *lick* in Spanish, paired with SPIT. These two pairs of words were problematic as NEs should have rated both pairs as unrelated to the English meaning (because they are unaware of the Spanish meaning); however, they rated them as related. Subsequently, these words were changed (PAN-flour to PAN-pita; LAME-spit to LAME-tongue). Participants were asked again a week later whether the new pairs were related or unrelated, the new pair of words were rated as unrelated. Once the changes were made, the Spanish survey was conducted. All bilingual participants rated the words as expected: the Spanish meaning of the interlingual homographs were rated as being related to the Spanish target words and unrelated to the English target words. All participants had the opportunity to identify any words that they did not know the meaning of during the study, none were found to be unfamiliar. These two surveys ensured that the prime – targets used for experiment 2a were appropriately related and unrelated to one another.

Design

A 2 (*Group*: NEs, Spanish-English bilinguals) x (2) (*Homograph type*: identical, near-identical) x (3) (*Target*: English related, Spanish related, unrelated) experimental design was used. Group is the between measure variables, and the repeated measures variables are homograph type and target.

Three versions of the task were created to counterbalance the stimuli material across conditions and to ensure that each participant only saw each interlingual homograph once in one of the experimental conditions (related to English meaning, Spanish meaning, or unrelated in meaning to either language). Each version of the experiment was split into three blocks with a total of: 72 experimental trials, 27 filler trials with word targets and 99 filler trials with non-word targets. Thus, participants made a total of 99 “yes” responses to real target words, and a total of 99 “no” responses to non-word targets. Presentations of trials were randomised for each participant.

Procedure

The entire experiment was conducted in the participants' L2 English and the computer task replicated the timings of Di Betta et al. (2015; see Appendix 43 for materials). The computer task was designed and presented on E-prime version 2.0 software (Psychology Software Tools, Pittsburgh, PA). Stimuli were presented in lower-case black letters (Courier New Font, size 16) on a silver background, except the prime, that was presented in uppercase. Each trial began with a 1000 ms blank screen, followed by a 100 ms fixation cross (+) and then the sentence. The sentence disappeared once the participant pressed the space bar on the keyboard (this indicated that the participant had finished reading the sentence), following this the prime appeared for 500 ms, and immediately after this, the target. Lexical decisions on the target were made by pressing the *x* (“no” response) and *m* (“yes” response) keys on a standard keyboard. The following trial would start automatically once the participant responded or after a time-out period of 3000 ms. For a visual depiction of the experimental sequence see Figure 8. Participants completed a series of 8 practice trials which were made up of sentences, primes, and targets which did not appear in the main experiment, the practice trials provided participants with feedback of whether the lexical decision was correct or incorrect. Participants were able to complete the practice trial multiple times.

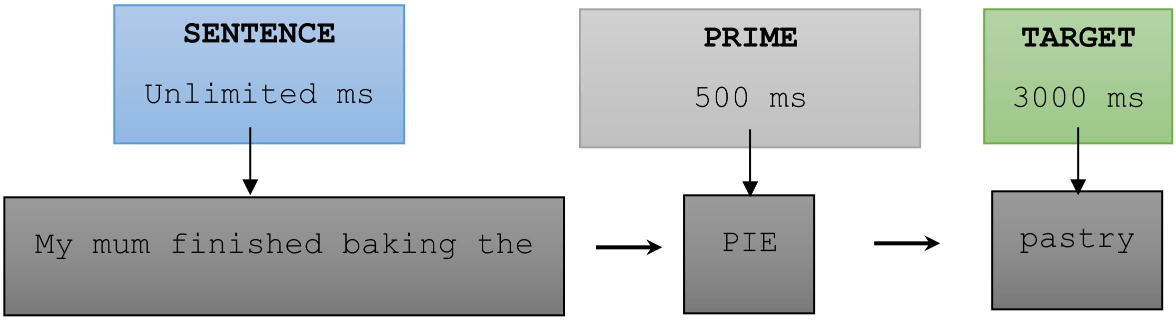

Figure 8.

Sequence flowchart of experiment (Study 2a).Before each trial begins a 1000 ms blank screen appears followed by a 100ms fixation cross (+). Prime is an example of an identical interlingual homograph, and the target of an English related target.

Additionally, participants responded to a total of 18 comprehension questions about the sentences read that required a yes- or no-response by pressing one of two response buttons on a standard keyboard. The questions were included in the experimental task as attention checks every 66 trials, to ensure that participants were indeed reading the sentences. After the task was performed, the participants completed a language background questionnaire, proficiency tests, and were asked to identify any experimental words they were unfamiliar with. Bilingual participants were given a £10 shopping voucher or course credits for their participation, NEs were given course credits for their participation after being debriefed. Overall, the experiment took approximately 60 minutes to complete.

Results

Four individuals were excluded from the analysis because they scored under 75% in accuracy on the attention questions. Moreover, one near-identical interlingual homograph was excluded from all analyses as it was found to be a cognate in Spanish (altar). Furthermore, an additional experimental near-identical interlingual homograph, one filler, and three non-words were excluded from the analysis because they were inadvertently repeated within the experimental blocks. There was therefore a total of 36 identical and 34 near-identical interlingual homographs included in the analysis, 26 trials with filler words, and 95 non-words. Filler words and non-words were not analysed.

Descriptive statistics in the form of mean RTs and mean errors percentages (%) for bilinguals and NEs are reported in Table 12. Bilinguals are slower compared to NEs and make more errors when making lexical decisions. NEs are faster and make fewer errors in the English related conditions compared to the Spanish related and unrelated target conditions suggesting priming may be present. Bilinguals are faster in the English related condition compared to the unrelated target condition suggesting priming may be present. However, they are particularly slower at reacting to Spanish related targets compared to unrelated targets in both identical (difference = 48 ms) and near-identical conditions (difference = 105 ms) compared to the unrelated which suggests possible interference of the L1 Spanish.

Table 12.

RTs (ms) and errors (%) descriptive statistics for identical and near-identical interlingual homographs.

	Group	**Identical English Related**	**Identical Spanish Related**	**Identical Unrelated**	**Near-Identical English Related**	**Near-Identical Spanish Related**	**Near-Identical Unrelated**
Mean RTs ms (*SD*)	Bilingual	828 (*206*)	898 (*291*)	850 (*209*)	791 (*176*)	921 (*302*)	816 (*186*)
	NEs	647 (*98*)	703 (*127*)	701 (*115*)	655 (*107*)	704 (*109*)	687 (*120*)
Mean Errors % (*SD*)	Bilingual	9% (.*11*)	13% (.*12*)	14% (.*13*)	5% (.*12*)	11% (.*09*)	8% (.*10*)
	NEs	3% (.*04*)	6% (.*07*)	8% (.*10*)	2% (.*05*)	7% (.*07*)	5% (.*08*)

Targets Collapsed	**Group**	**English related**	**Spanish Related**	**Unrelated**
Mean RTs ms (*SD*)	Bilingual	810 (*183*)	909 (*290*)	833 (*191*)
	NEs	651 (*99*)	704 (*113*)	694 (*108*)

Mean Errors % (*SD*)	Bilingual	9% (*.08*)	10% (*.08*)	12% (*.07*)
	NEs	5% (*.05*)	7% (*.06*)	7% (*.07*)

Note – Identical and near-identical interlingual homographs have been collapsed in the table to form three target conditions: English related, Spanish related and unrelated targets to provide a whole picture of interlingual homograph effects.

For both RTs and errors, two analyses of variance (ANOVAs) are performed; one analysis is performed with *participants* (F_1) and the other with *items* (F_2) as the random factor. For the analysis of RTs, only correct responses were considered. The F_1 design is a 2 x (2) x (3) and involves analysis of the group (NEs, Spanish-English bilinguals) as a between-subject variable, and homograph type (identical, near-identical) and target (English related, Spanish related, unrelated) as the repeated measures variables. In the F_2 analyses, it is a 2 x (2) x (3) design, with homograph type acting as the between-subject variable; group and target as the repeated measures variables. In the few cases in which the $F2$ analyses do not coincide with the $F1$ analysis, theoretical interpretation of the statistic will be based on the $F1$ analysis. Subsequently, any such interpretations should, therefore, be treated with caution.

Reaction Time Analysis

The incorrect responses (6.16% of the data), trials where the sentence reading time was below 500 ms (0.5% of the data), and the trials with RTs below 200 ms and exceeding a criterion of +/-2.5 SD for an individual participant's mean (3.12% of the data) were excluded from the analysis. To ensure that participants had read the sentences and subsequently semantic priming had occurred, any homograph meanings that were unknown to the participant were also excluded from the analysis (1.28% of the data).

The analysis of RTs revealed a main effect of group F_1 (1, 58) = 16.59, $p < .001$, $\eta_p2 = .222$; F_2 (1, 68) = 466.89, $p <.001$, $\eta_p2=.873$. Bilinguals took significantly longer to respond than the NEs (850 ms vs. 683 ms, respectively). There was no main effect of homograph type (F_1 (1, 58) = 1.12, $p =.295$, $\eta_p2 = .019$; F_2 (1, 68) = .65, $p = .423$, $\eta_p2 = .009$). There was a main effect of target, F_1 (1.46, 84.53) = 19.06, $p <.001$, *Greenhouse-Geisser corrected,* $\eta_p2 = .247$; F_2 (2, 136) = 10.74, $p <.001$, $\eta_p2 = .136$. Pair-wise comparisons indicated that overall, RTs were significantly faster in the English related target condition (731 ms) compared to unrelated targets (763ms, p =.001, *Bonferroni corrected*) suggesting that overall priming was present. Spanish related targets were significantly slower in comparison to unrelated targets (p = .003, *Bonferroni corrected*).

There was a significant interaction between target and group, F_1 (2, 116) = 3.91, p =.023, η_p2= .063; F_2 (2, 136) = 5.02, p =.008, η_p2= .069. There were no significant interactions between homograph type and group, homograph type and target, and no three-way interaction between homograph type, target, and group (see Appendix 44).

To unpick the target and group interaction, two repeated measures ANOVAS were conducted for each group (see Appendix 45). Identical and near-identical homograph types were collapsed together into three target conditions because there were no significant main effect or interactions relating to this variable. See Figure 9 for a visual depiction of the priming differences between unrelated targets and the target (English related, Spanish related) RTs.

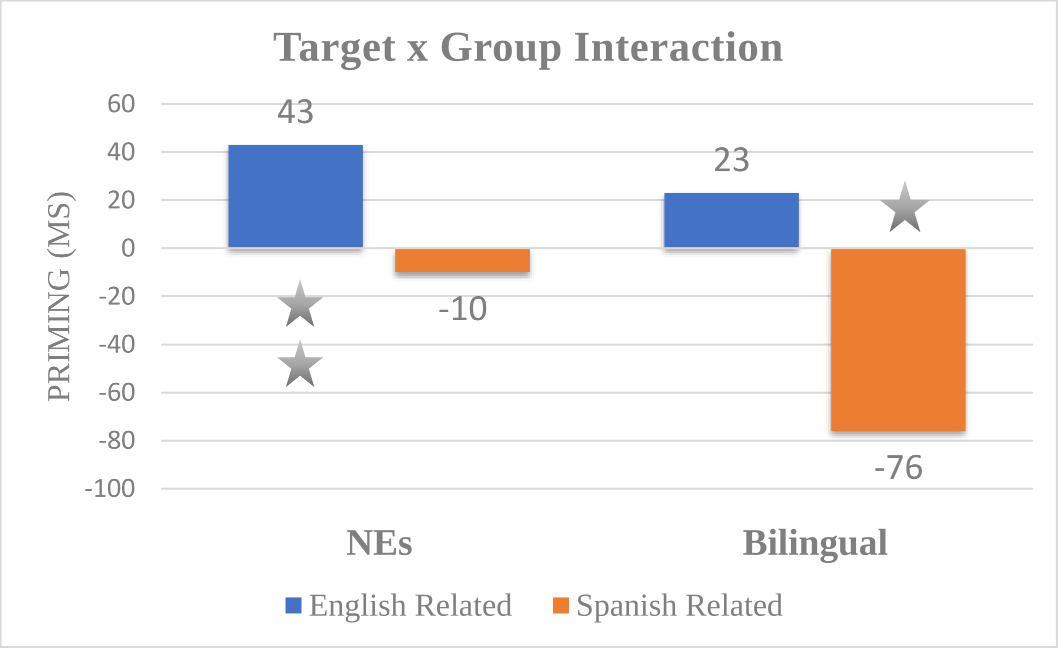

Figure 9.

Priming differences (ms) between unrelated target (baseline) RTs minus the target RTs.

In Figure 9 the X axis shows the two groups: NEs and bilinguals, and each condition. The blue is the English related condition minus the unrelated condition; and the orange is the Spanish related condition minus the unrelated condition. One star reflects significance at p = .05, two stars reflect significance at p < .001.

For the *NEs group*, there was a large significant main effect of target, F_1 (2, 70) = 19.21, p <.001, η_p2 = .354; F_2 (2, 136) = 7.93, p =.001, η_p2= .104). As expected, pair-wise comparisons revealed that RTs were significantly faster for English targets (651 ms) compared to unrelated targets (694 ms; p <.001, *Bonferroni corrected*) suggesting priming was present. There was no significant difference between unrelated and Spanish related targets (704 ms*;* p = .756, *Bonferroni corrected*). RTs were also significantly faster for English targets compared to Spanish related targets (p <.001, *Bonferroni corrected*) which is expected as NEs were not familiar with the Spanish language and therefore would see it as an unrelated target.

In the *bilingual group*, there was a large main effect of target, F_1 (1.26, 29.04) = 7.38, p =.007, *Greenhouse-Geisser corrected,* η_p2 = .243; F_2 (2, 136) = 9.71, p <.001, η_p2= .125. Pair-wise comparisons revealed that there were no significant differences between English related targets (810 ms) compared to unrelated targets (833 ms; p = .452, *Bonferroni corrected).* Spanish related targets (909 ms) were responded significantly more slowly compared to unrelated targets (p = .032, *Bonferroni corrected),* suggesting interference from the bilinguals' L1. The targets were related to the Spanish meaning of the prime (i.e., homograph), which may suggest there was a semantic interference cost for tapping into the Spanish language.

Error Analysis

The analysis of error data revealed a large main effect of group F_1 (1, 58) = 8.49, $p = .005$, $\eta_p2 = .128$; F_2 (1, 68) = 31.89, $p <.001$, $\eta_p2 = .319$. Bilinguals produced more errors (10.1%) than the NEs (5.1%). There was a main effect of homograph type, F_1 (1, 58) = 14.28, $p < .001$, $\eta_p2 = .198$; F_2 (1, 68) = 2.47, $p = .121$, $\eta_p2 = .035$. More errors were produced in the identical condition (8.9%) compared to the near-identical (6.3%). There was a main effect of target, F_1 (1.81, 104.86) = 12.40, $p < .001$, *Greenhouse-Geisser corrected,* η_p2= .176; F_2 (1.81, 123.32) = 2.94, $p =.062$, *Greenhouse-Geisser corrected,* $\eta_p2 = .041$. There were significantly fewer errors made in the English related condition (4.8%) compared to the unrelated target conditions (8.8%; $p = .002$), whilst Spanish related targets (9.2%) and unrelated targets did not significantly differ in the number of errors made ($p = 1$). There were also significantly fewer errors made in the English related condition compared to the Spanish target condition ($p < .001$). However, the homograph type and target effects should be treated with caution and may not be reliable due the F_2 not reaching significance, and thus being inconsistent with the F_1 analysis.

There was a significant interaction between homograph type and group, F_1 (1, 58) = 4.20, $p = .044$, $\eta_p2 = .068$; F_2 (1, 68) = 2.61, $p = .111$, $\eta_p2 = .037$. There was a significant interaction of homograph type and target, F_1 (2, 116) = 3.46, $p = .035$, $\eta_p2 = .056$; F_2 (2, 136) = .67, $p = .513$, $\eta_p2 = .010$. There were no significant interactions between target and group, and no three-way interaction between homograph type, target, and group (see Appendix 46).

While the participant and item analysis are not both consistently significant in the homograph type and group, and homograph type and target interactions, these interactions will still be unpicked because of the significant F^1 analysis, but results should be treated with caution. A repeated measures ANOVA (2: identical, near-identical x 3: English related, Spanish related, unrelated) was conducted for each group: NEs and bilinguals (see Appendix 47).

For the *NEs group,* there was no main effect of interlingual homograph type, F_1 (1, 35) = 2.21, p =.146, $\eta_p 2$ = .059; F_2 (1, 68) = .83, p =.366, $\eta_p 2$ = .012. There was a large main effect of target, F_1 (2, 70) = 7.16, p =.001, $\eta_p 2$ = .170; F_2 (1.73, 117.35) = 3.64, p =.035, *Greenhouse-Geisser corrected,* $\eta_p 2$ = .051. As expected, pair-wise comparisons revealed that there were significantly fewer errors (2.6%) in the English related targets compared to the unrelated targets (6.6%; p = .008, *Bonferroni corrected*). There were no differences between unrelated targets and Spanish related targets (6.2%; p = 1, *Bonferroni corrected*). There were also significantly fewer errors in the English related targets compared to the Spanish related targets (p = .005, *Bonferroni corrected*), which is expected as NEs were not familiar with the Spanish language and therefore would see it as an unrelated target. There was no significant interaction between homograph type and target.

For the *bilingual group,* there was a main effect of interlingual homograph type, F_1 (1, 23) = 11.42, p =.003, $\eta_p 2$ = .332; F_2 (1, 68) = 3.01, p =.088, $\eta_p 2$ = .042. Bilinguals made more errors when primed by identical interlingual homographs (12%) compared to near-identical homographs (8.1%; p =.003). There was a large main effect of target, F_1 (1.55, 35.63) = 5.58, p =.007, *Greenhouse-Geisser corrected,* $\eta_p 2$= .195; F_2 (1.83, 124.34) = 1.87, p =.162, *Greenhouse-Geisser corrected,* $\eta_p 2$ = .027. There were no significant differences in errors between English related targets (7%) and unrelated targets (11%; *p* =.189, *Bonferroni corrected*) suggesting no priming was present, or between Spanish related (12.3%) and unrelated targets (*p* = .965). Though, bilinguals did significantly make fewer errors in the English related targets compared to the Spanish related targets (*p* = .006). There was no significant interaction between homograph type and target.

Discussion

The present study set out to explore whether bilinguals' activation of their L1 is modulated by sentence meaning in an L2 task and whether there are processing differences between interlingual homographs with differing amounts of cross-linguistic orthographic overlap. In doing so, it aimed to conceptually replicate Di Betta et al.'s (2015) work with Polish-English bilinguals, to see whether similar findings could be replicated in Spanish – English bilinguals using the same task paradigm.

The first aim set out to explore to what extent the participant's L1 Spanish is activated during an L2 task, and whether the L1 is modulated by sentence meaning. For the NEs who were used as the control group, it was predicted that they would show significant priming to English related targets, and no differences between the unrelated and Spanish related targets conditions as they had no experience of the Spanish language and so prime and target word pairs would be unrelated to them. NEs behaved as expected, demonstrating priming for the English related targets in the RTs, and no differences between unrelated and Spanish related targets. Moreover, fewer errors were made in the English related targets compared to the other target conditions; and there were no differences between unrelated and Spanish related targets.

For the bilinguals, it was predicted, as in Di Betta et a's study (2015), that one would see interference from the L1. Overall bilinguals took longer to respond and made more errors compared to NEs control group in all target conditions, which is to be expected as the task is in their L2. This pattern of results is consistent with literature showing that bilinguals take longer and make more errors when conducting the task in their L2 compared to their L1 (e.g., Lemhöfer & Dijkstra, 2004). Furthermore, the results showed that in both the identical and near-identical interlingual homograph there was a trend towards priming for the English related targets, though this did not reach statistical significance. However, when the two types of interlingual homographs were collapsed, there was significant negative priming for the Spanish-related targets, which suggests that their L1 Spanish has been activated and is slowing down the RT to the Spanish-related target compared to the unrelated one.

Overall, bilinguals made more errors than NEs; and more errors were produced in the identical condition compared to the near-identical interlingual homograph condition. In the identical interlingual homographs the orthography of the prime is the same for L1 Spanish as the bilinguals' L2 English; therefore, this may be confusing when making lexical decisions to targets related to these primes. This would be particularly confusing as models such as the BIA+ model (Dijkstra & van Heuven, 2002) suggest that these words may overlap in their lexical representations in the lexicon and are strongly connected with each other.

In this study the longer RTs in the Spanish related targets (non-target meaning) compared to the unrelated targets could be explained by the BIA+ (Dijkstra & van Heuven, 2002) as a disproportionately strong effect of lateral inhibition; because the lexical representations of the interlingual homographs are identical (or very similar), this competition is stronger compared to control words. The IC model (Green, 1998; 2003) can also explain why inhibitory effects are present. The language task schemes in the IC model (Green, 1998) are moderated by a supervisory attentional system that regulates their activity by inhibiting the task schema for the non-target language, it can limit interference and maintain activation for the target language. Therefore, the supervisory attentional system has to work harder to maintain the activation of the target language while inhibiting the non-target language. Thus, both models suggest that the initial conflict between two languages in comprehension is resolved by a mechanism of active inhibition. In this study, the activation of the non-target L1 Spanish is strong (as suggested by the negative priming), however, L1 activation needs to be inhibited while also maintaining activation of the L2 meaning to meet the task demands of being able to make a lexical decision.

The inhibitory effects in this study seen in the bilinguals are consistent with previous studies interpretations that found that bilinguals are slower when processing interlingual homographs compared to control words (e.g., De Groot et al., 2000; Dijkstra et al., 2000; Dijkstra et al., 1998; Libben & Titone, 2009; Pivneva et al., 2014; Poort et al., 2016; Von Studnitz & Green, 2002). These results are also consistent with previous literature that has found inhibitory effects with interlingual homographs for tasks which are conducted in a participants' L2. For example, Lemhöfer and Dijkstra (2004) conducted variants of an LDT and gave participants 1500 ms to respond to the interlingual homographs; the first experiment included both L1 Dutch and L2 English words, the other only L2 English words. For the Dutch-English LDT facilitatory effects were found, but when the task was swapped to solely English, inhibitory effects were found. Inhibition was interpreted as the L1 non-target language interfering with the task. However, the current study used interlingual homographs as primes and not targets, and only used written L2 words as targets, and therefore bilinguals needed to rely on the meaning of the interlingual homograph and not just the orthographic representation. Therefore, the present experiment extends Lemhöfer and Dijkstra's (2004) findings and shows that in a written priming sentence context the same inhibitory effects can be present even when an interlingual homograph is used a prime and not just a target. Our task only used English target words and inhibitory effects are consistent with those of Lemhöfer and Dijkstra (2004) LDT variant (experiment 1) that only included L2 English targets. As we saw above, Lemhöfer and Dijkstra (2004) also conducted a variant of an LDT where L1 targets were introduced, and this resulted in positive priming. Subsequently, a potential future avenue of research could be to introduce Spanish words as targets in the current study's experimental paradigm to see whether including targets of both languages would have a positive or negative effect on priming. This would test the interpretation that the L1 was activated in the current study, if the L1 is activated non-selectively then including stimuli from the L1 Spanish as targets might make the negative priming effects stronger in the Spanish related target condition. This would be able to provide more evidence to the notion that both languages are activated during bilingual language processing.

A further aim of the study was to identify whether there were processing differences for interlingual homograph types with a varying degree of orthographic overlap. For the NEs it was predicted that there would be no differences as the stimuli only belonged to one lexicon (English) and this was the case, participants did not significantly differ in RTs and accuracy in Spanish related targets compared to unrelated which was interpreted in that both targets were processed similarly. In contrast, for the bilingual group it was predicted that there would be differences in the priming effects elicited by the interlingual homographs, with potentially facilitatory effects for near-identical and inhibitory effects for identical homographs (Di Betta et al., 2015). Furthermore, based on the cognate literature where facilitation effects are reduced as orthographic overlap is increased (e.g., van Assche et al., 2011; Dijkstra et al., 2010; Peeters et al., 2013), it was predicted that there would be stronger L1 effects present for targets primed by identical interlingual homographs compare to near-identical ones. The RT results suggest that identical and near-identical interlingual homographs did not significantly differ from one another as both homograph types elicited negative priming for the Spanish-related targets. Bilinguals did, however, make significantly more errors to targets that were primed by the identical interlingual homographs compared to the near-identical interlingual homographs. The results are consistent with the cognate literature in that the same effect is present for both identical and near-identical interlingual homographs, however, the overall magnitude of the inhibitory effects is not moderated by orthographic overlap as the near-identical interlingual homographs were only 22 ms slower compared to the identical condition and this was a non-significant difference.

Moreover, while the identical interlingual homographs inhibitory effects are consistent with Di Betta et al.’s (2015) findings, the near-identical inhibitory effect is not. Di Betta et al. (2015) found processing differences of Polish-related targets according to homograph type; significant priming for near-identical interlingual homographs and negative priming effects for identical interlingual homographs. Though the task paradigm was identical, there were some key differences with the present study. Firstly, it is not clear whether the bilinguals used in Di Betta et al.’s (2015) study were as proficient in their L2 in comparison to those in the present experiment who were proficient late bilinguals and whose proficiency in L2 was assessed in a more objective manner (the participants from the Di Betta et al's (2015) study were only asked to self-rate their language proficiency in English). According to the IC model (Green, 1998) the amount of inhibition used to resolve interference of the non-target language should be proportional to the level of activation of a given language. Thus, the higher the basic level of the non-target language activation, the more inhibition is required to supress it; the strength of inhibition is a function of L2 proficiency. For example, when a bilingual uses their weaker L2, word candidates from the more dominant L1 are strongly activated and compete with the target language candidates and therefore need to be strongly inhibited (Green, 1998). Subsequently, the more strongly a representation has been inhibited, the longer time it requires to overcome inhibition which may have contributed to the inhibitory effects seen in both identical and near-identical interlingual homographs. Evidence for this claim could be provided by manipulating the proficiency of bilinguals and using this to explore whether negative priming would be consistent across orthographic overlap and proficiency.

The second difference between the current study and Di Betta et al. (2015) is that two different languages were used, and the difference in effects may be due to orthographic and phonological differences between the Spanish and Polish language on one side, and English on the other. For example, speech rhythm in languages is based on the observation that different languages give rise to different types of prosodies. The English language is stress-timed, Spanish syllable-timed, but Polish seems to be different from any other language studies and may constitute a new rhythm class (Ramus, Dupoux, & Mehler, 2003). Therefore, the rhythm of the languages may be affecting the way bilinguals process interlingual homographs.

Furthermore, the third difference is the Levenshtein distances of the near-identical interlingual homographs between the current study and Di Betta et al. (2015) are different. In the current study the average distance between homographs was 1.2, but in Di Betta et al.'s (2015) experiment the average is 3 (personal communication). Thus, it could be that the difference between the experiments' results is due to the extent of the cross-linguistic orthographic overlap; however, this is an unavoidable confound as the distances between interlingual homographs are constrained by the characteristics of the language. It is also unavoidable that the linguistic properties of the interlingual homographs such as word frequency will also be different across languages, hence this could be influencing the difference in results. Nevertheless, in this experiment, the bilinguals’ L1 is may be activated more strongly as the lexical overlap between the near-identical interlingual homographs is greater compared to Polish – English near-identical interlingual homographs. The L1 Polish meaning is being activated when bilinguals process targets primed by the near-identical interlingual homographs as priming is present, however, because there is not much lexical overlap cross-linguistically, the Polish reading of the interlingual homograph is not as strongly inhibited by the English reading of it as would be the case for any semantically related word pairs in a priming task. This brings to the table another interesting issue, which is that most of the literature on bilingualism focuses on Dutch or Spanish as the participants' L1. The inconsistency between our results and those of Di Betta et al (2015) calls for a greater variety of L1s to be explored in order to paint a more complete picture of all the characteristics at play in bilingual language processing. It is interesting to note that, for instance, when Polish-English identical interlingual homographs are used in a semantic task such as the Semantic Relatedness Judgement task, results are similar to those reported for Spanish-English interlingual homographs (Durlik et al., 2016). Therefore, the task used is important too (for brief discussion of task demands in interlingual homographs see Poort & Rodd, 2019).

To conclude, our data demonstrates that even in an exclusively L2 context bilinguals activate both L1 and L2 meanings of identical and near-identical interlingual homographs, that is, a sentence context is not enough in inhibiting completely the irrelevant L1 Spanish meaning. The results are consistent with the results from study 1b in the previous chapter and the theory that inhibitory processes are required to resolve the interference and inhibit the irrelevant L1 meaning, and this pattern is consistent in both the identical and near-identical interlingual homographs.

Study 2 (b)

Study 2a demonstrated that NEs only activated the English meaning of the interlingual homographs, as expected, whilst Spanish – English bilinguals activated both the L1 and L2 meanings of the interlingual homographs. This was inferred because of the significant negative priming that was present for the Spanish-related meaning of the interlingual homographs, and this finding was interpreted as evidence of inhibition of the non-relevant L1 reading of the interlingual homographs. Experiment 2a involved highly proficient bilinguals and a prime duration (that is, the time the interlingual homograph was presented before the target appeared on screen) of 500 ms. However, it is unclear whether the L1 meaning activation from the Spanish related targets (i.e., negative priming) is due to interference or inhibition. Therefore, to shed some light on this, the goal of Study 2b is to unpick whether the negative priming seen for the Spanish related targets is due to interference or inhibition by using the exact paradigm as study 2a but manipulate the prime duration from 500 ms to 200 ms.

There is evidence from both monolingual (Gernsbacher, Varner, & Faust, 1990; Elston-Güttler & Friederici, 2005) and bilingual studies (Beauvillain & Grainger, 1987; Martín et al., 2010) shows that it is possible to reveal the activation of an ambiguous word's different meanings by probing them at different points in time. For instance, Elston-Güttler and Friederici (2005) used sentences to contextually prime the dominant or subordinate meanings of homonyms in native NEs and proficient German (L1) – English (L2) bilinguals. Homonyms are words that are identical in spelling but have different semantic meanings within a language, such as "bark" semantically associated to a "dog" or "tree". When the homonym prime was onscreen for 200 ms, natives and non-natives showed priming for both its meanings whether contextually appropriate or not. However, when the prime duration was altered to 500 ms, both groups showed priming effects for *only* the contextually appropriate targets. However, this study also measured event-related potentials (ERPs) and while generally the ERPs were consistent with the behavioural data, in the 500 ms condition non-natives showed activation of the non-target meaning despite the behavioural data suggesting that only the target meaning was activated.

Therefore, from previous literature we know that native speakers have settled on the correct interpretation of the homograph when the prime is presented for 500 ms (e.g., Elston-Güttler & Friederici, 2005; Di Betta et al., 2015). For instance, the Polish bilinguals in Di Betta et al. (2015) still showed effects of L1 activation with a prime presentation of 500 ms. Furthermore, in light of the results of Study 2a, one could suggest that shortening the length of the prime presentation may allow one to bring to light stronger L1 effects (that is a situation where the participant has not settled on the correct meaning of the ambiguous word) as in the case of Elston-Güttler and Friederici (2005). L1 effects during this shortened prime duration may support the idea of interference from the L1, and this could be facilitatory (positive) priming effect which may indicate that the L1 meaning has not been inhibited yet, or inhibitory (negative) priming effect as we saw in Study 2a. Whereas no effects of the L1 may point towards the argument that inhibition was used in study 2a (i.e., 200 ms prime duration may be too short to activate and then inhibit the irrelevant meaning).

Method

Participants

A total of twenty-four Spanish-English bilinguals (14 females and 10 males; *mean age* = 26 years, *SD* = 4.05) from the Sheffield and Swansea area and thirty NEs (7 men; *mean age* = 19 years, *SD* = 4.31) from Sheffield Hallam University participated in this experiment. None of them had taken part in Experiment 2a. The participant's characteristics are reported in Table 13.

All bilinguals were educated to at least A-levels or equivalent level; nine had graduated with a bachelor's degree, and eleven had a master's degree. In addition to English and Spanish, thirteen bilinguals had experience with another language, one with more than one language. All bilinguals had received a score of 6.5 or higher in the IELTS.

All NEs were educated to at least A-levels or equivalent, with one participant with a master's degree. Eight NEs had no experience with another language; twenty-one had experience with one other language, and one NE had experience with more than one language. The same inclusion criteria used for Experiment 2a were applied and NEs were required to *not* be actively learning an L2 and to have no prior knowledge of the Spanish language. No participants reported a language disability or neurological disorders.

Table 13.

Participant language characteristics.

	Bilingual	NEs	t-test (*p*)
Age in years (*SD*)	26 (*4.1*)	19 (*4.3*)	< .001
LexTALE	83.8 (*6.3*)	92.3 (*5.2*)	< .001
LexTALE-SPA	96.9 (*1.9*)	NA	NA
Living in L2 speaking countries (months)	27.2 (*27.1*)	All their life	NA
Self-rated proficiency English[1]	3.4 (.*5*)	4.0 (.*1*)	< .001
Self-rated proficiency Spanish	4.0 (.*4*)	NA	NA

Daily use English[2]	4.9 (*.4*)	5	ns
Daily use Spanish	4.92 (*.2*)	NA	NA

Note. NA = not applicable, ns = non-significant. Standard deviation is reported in brackets unless stated otherwise. [1]Self-ratings ranged from 0 to 4 (0 = not at all, 1 = not well, 2 = I can, but with a lot of difficulty, 3 = well, but with a little difficulty, 4 = very well) and were rated on reading, speaking, and understanding; averages are shown. [2]Self-ratings of daily use ranged from 1 to 5 (1 = rarely, 2 = 1-2 times per month, 3 = once a week, 4 = 2-3 times a week, 5 = every day) and were rated on reading, speaking, and understanding; averages are shown. Some of the participants reported some knowledge and use of L3. Independent t-tests were conducted to compare Bilingual and NEs conditions in the language characteristics and can be seen in the third column.

Stimulus, Design and Procedure

The stimuli, design, and procedure were identical to that of Study 2(a) except that the duration of the prime was altered to 200 ms.

Results

The incorrect responses (9.14% of the data), trials where sentence reading time was below 500ms (0.70% of the data), and those with RTs below 200 ms and exceeding a criterion of +/-2.5 SD for an individual participant's mean (2.84% of the data) were excluded from the analysis. Any homograph meaning that was unknown to the participant resulted in the interlingual homograph being also excluded from the analysis (0.89% of the data).

The mean RTs and mean error percentages (%) for bilinguals and NEs are reported in Table 14. Bilinguals were slower compared to NEs and make more errors when making lexical decisions. NEs were faster and make fewer errors in the English related conditions compared to the Spanish related and unrelated target conditions suggesting priming is present in both prime-time durations. Bilinguals were faster in the English related conditions compared to the Spanish related and unrelated target conditions again suggesting priming is present. There does not appear to the interference of the L1 Spanish in the identical interlingual homograph Spanish target condition, as the Spanish target condition and unrelated target condition are similar in RTs (difference = 13 ms). Whereas, in the near-identical interlingual homograph primed Spanish related target condition there appears to be some interference of the L1 Spanish as bilinguals were 42 ms slower compared to the unrelated target condition. However, the inferential statistics will test these interpretations from the descriptive statistics.

Table 14.

RTs (ms) and error (%) descriptive statistics for identical and near-identical interlingual homographs at 200 ms.

	Group	**Identical English Related**	**Identical Spanish Related**	**Identical Unrelated**	**Near-Identical English Related**	**Near-Identical Spanish Related**	**Near-Identical Unrelated**
Mean RTs ms (*SD*)	Bilingual	888 (*192*)	999 (*203*)	1012 (*229*)	972 (*227*)	1060 (*222*)	1018 (*215*)
	NEs	746 (*157*)	773 (*189*)	791 (*175*)	738 (*174*)	779 (*170*)	782 (*199*)
Mean Error % (*SD*)	Bilingual	10% (*.09*)	7% (*.09*)	13% (*.10*)	8% (*.10*)	14% (*.13*)	10% (*.09*)
	NEs	6% (*.08*)	9% (*.09*)	10% (*.11*)	3% (*.05*)	7% (*10*)	6% (*.08*)

Targets Collapsed	**Group**	**English related**	**Spanish Related**	**Unrelated**
Mean RTs ms (*SD*)	Bilingual	930 (*202*)	1030 (*203*)	1015 (*211*)

	NEs	742 (*158*)	776 (*173*)	787 (*179*)
Mean Error % (*SD*)	Bilingual	7% (.*10*)	13% (.*10*)	11% (.*09*)
	NEs	3% (.*03*)	6% (.*06*)	7% (.*09*)

Note - Identical and near-identical interlingual homographs have been collapsed in the table to form three target conditions: English related, Spanish related and unrelated targets.

Reaction Time Analysis

The analysis of RTs revealed a main effect of group F_1 (1, 52) = 20.28, p <.001, η_p2 = .281; F_2 (1, 67) = 381.07, p < .001, η_p2 = .850. Bilinguals took significantly longer to respond than the NEs (992 ms vs 768 ms, respectively). There was a main effect of homograph type, F_1 (1, 52) = 7.18, p =.010, η_p2 = .121; F_2 (1, 67) = 2.25, p < .138, η_p2 = .033. RTs were significantly faster in response to targets that followed identical interlingual homographs (868 ms) compared to those following near-identical interlingual homographs (891 ms). There was a main effect of Target, F_1 (2, 104) = 26.43, p <.001, η_p2 = .337; F_2 (2, 134) = 9.97, p <.001, η_p2 = .129. Pair-wise comparisons indicated that RTs were significantly faster in the English related target condition (836 ms) compared to the unrelated targets (901 ms, p <.001, *Bonferroni corrected*) suggesting that overall, priming was present. There were no significant differences between Spanish related targets (903 ms) and unrelated targets (p = 1). English related targets were also significantly faster than Spanish related targets (p < .001, *Bonferroni corrected*). However, the homograph type effect should be treated with caution and may not be reliable due the F_2 not reaching significance, and thus being inconsistent with the F_1 analysis. Any interpretations that are made in this section will be based on the F_1 analysis.

There was a significant interaction between interlingual homograph type and group, F_1 (1, 52) = 9.74, p =.003, η_p2 = .089; F_2 (1, 67) = 3.90, p =.053, η_p2 = .055. There were no significant differences for NEs for targets that were primed by identical (770 ms) or near-identical (766 ms) interlingual homographs (t (29) = .35, p = .732), and this is interpreted as NEs processing both homograph types similarly. Whereas bilinguals were significantly slower making lexical decisions to targets primed by near-identical interlingual homographs (1015 ms) compared to identical (966 ms) interlingual homographs (*t* (23) = 3.71, *p* = .001). There was also a significant interaction between target and group, F_1 (2, 104) = 4.95, *p* =.009, $\eta_p 2$ = .087; F_2 (2, 134) = 2.10, *p* =.127, $\eta_p 2$ = .030. No interaction was found between homograph type and target, and there was no three- way interaction between homograph type, target, and group (see Appendix 48).

To unpick the interaction between target and group, two repeated measures ANOVAS were conducted, one for each group: NEs and bilinguals (see Appendix 49). Identical and near-identical interlingual homographs were collapsed to make three target conditions: English related, Spanish related, and unrelated. See Figure 10 for visual depiction of priming effects.

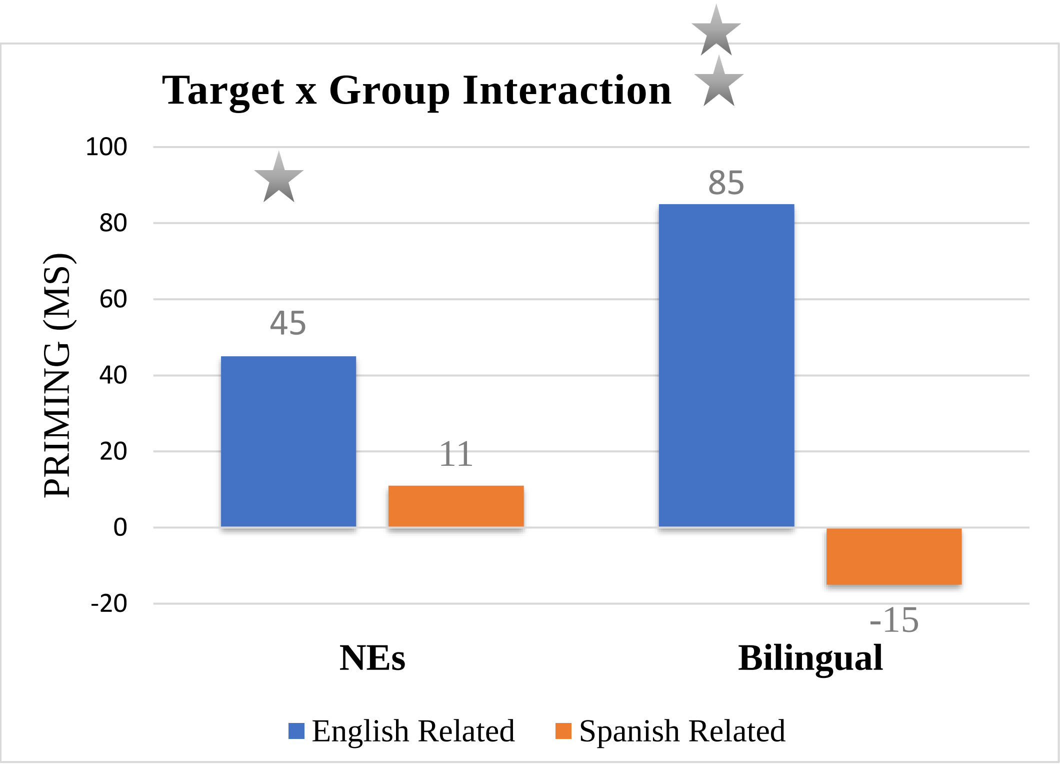

Figure 10.

Priming differences (ms) between unrelated target (baseline) RTs minus the target RTs.

In Figure 10, the X axis shows the two groups: NEs and bilinguals, and each condition. The y axis reports the priming as difference between unrelated and related condition in ms (English or Spanish). The blue refers to the English related condition the orange to the Spanish related condition. One star reflects significance at $p = .05$, two stars reflect significance at $p < .001$.

In the *NEs* group, the analysis revealed a large main effect of target, F_1 (2, 58) = 8.73, $p <.001$, $\eta_p2 = .231$; F_2 (2, 136) = 4.25, $p =.016$, $\eta_p2 = .059$. Pair-wise comparisons revealed significantly faster RTs in the English related targets (742 ms) compared to the unrelated targets (787 ms; $p =.005$, *Bonferroni corrected*). English related targets were also faster compared to Spanish related targets (776 ms; $p =.017$, *Bonferroni corrected*). Thus, overall priming was still present for the NEs group. As expected, there were no significant differences between Spanish related targets and unrelated targets ($p = .711$, *Bonferroni corrected*).

In the *bilingual group*, the analysis revealed a large main effect of target, F_1 (2, 46) = 16.42, $p <.001$, $\eta_p2 = .417$; F_2 (2, 134) = 7.09, $p =.001$, $\eta_p2 = .096$. Pair-wise comparisons revealed significantly faster RTs to the English related targets (930 ms) compared to the unrelated targets (1015 ms; $p <.001$, *Bonferroni corrected*) suggesting that this time priming for the English meaning reached statistical significance. However, there were no significant differences between Spanish related targets and unrelated targets ($p = 1$, *Bonferroni corrected*).

To summarise, bilinguals took longer to respond than NEs, which is to be expected as the task is in their L2. NEs behaved as expected, demonstrating priming for the English but not the Spanish meaning, which is unknown to them. Priming was also present in the bilingual group for the English meaning, suggesting that these participants were activating the meaning of the interlingual homographs relevant to the sentence context. However, there was no significant priming of the Spanish meaning. Numerically, though, the negative priming evident in Experiment 2a appeared reduced in the current experiment.

Error Analysis

There was no main effect of group, F_1 (1, 52) = 3.75, p = .058, η_p2 = .128; F_2 (1, 67) = 7.14, p =.009, η_p2 =.096. There was a main effect of homograph type, F_1 (1, 52) = 11.66, p = .001, η_p2 = .183; F_2 (1, 67) = .93, p =.338, η_p2 = .014, with more errors made on target following identical interlingual homographs (10%) compared to those following the near-identical ones (7%). There was no main effect of target, F_1 (2, 104) = .45, p = .638, η_p2 = .009; F_2 (2, 134) = 2.08, p =.129, η_p2 = .030.

There was a significant interaction of homograph type and target, F_1 (2, 104) = 5.29, p = .007, η_p2 = .092; F_2 (2, 134) = 1.86, p = .159, η_p2 = .027. There were no interactions between homograph type and group; target and group; and there was no significant three-way interaction between homograph type, target, and group (see Appendix 50). However, since this indicates that language group did not behave significantly different from each other in the different conditions, no further analyses were carried out.

To summarise, there were no differences in errors between groups, and target. However, more errors were made to targets primed by identical homographs compared to near-identical ones.

Discussion

Study 2b aimed to investigate if the activation of bilinguals' L1 Spanish in Study 2a was due to inhibition or interference and could be uncovered by modifying the length of time the prime was presented for (200 ms vs 500 ms in Experiment 2a). It further aimed to explore whether there are processing differences between interlingual homographs with differing amounts of cross-linguistic orthographic overlap.

Overall, the results showed that bilinguals took longer to respond compared to NEs in all target conditions, which, as mentioned before, is consistent with the results from previous literature (e.g., Lemhöfer & Dijkstra, 2004). For the NEs, findings confirmed the prediction that significant priming would only be evident with English-related targets and that there would not be differences between the unrelated and Spanish related targets conditions as these participants had no experience of the Spanish language. For the bilinguals, priming was present for the English meaning of the interlingual homograph, which indicates that the meaning of the interlingual homograph relevant to the sentence context had been activated when this shorter prime duration was used. However, the experiment did not find evidence of direct interference from the L1 (which would have been shown by finding significant priming for the Spanish related meaning) at this prime duration. This finding therefore, might suggest that the negative priming seen in study 2a might have been due to inhibition of the L1 meaning; however, it is still not clear whether it was due to interference or inhibition of the L1 as the experiment does not allow one to identify where the interference is coming from.

These results are consistent with previous priming tasks that have not found any effects of interlingual homograph processing (e.g., De Bot et al., 1995; Dijkstra et al., 1998; Gerard & Scarborough, 1989). Dijkstra et al. (1999) suggested that these studies did not find any effects because they failed to control for phonological similarity. The positive influence of orthographic overlap and the negative influence of phonological overlap generates an overall a null overall effect on RTs (Dijkstra et al. 1999). However, in this study phonological similarity was controlled for, and all identical and near-identical interlingual homographs have distinct pronunciation. The difference may lay in the task used since, the cited studies used isolated paradigms and this study uses a sentence context and differences in the prime duration used. In Study 2a we saw that it is more common for studies using interlingual homographs in sentence and priming paradigms to find longer RTs when processing these types of stimuli (e.g., De Groot et al., 2000; Libben & Titone, 2009; Pivneva et al., 2014; Poort et al., 2016; Titone et al., 2011; Von Studnitz & Green, 2002). Furthermore, these studies focused on participants responding to the interlingual homographs, and not to related targets such as in this experiment, so it may be simply due to different task demands. Though, Study 2a did find negative priming, so it may be due to the prime duration used; to test this notion, the next section compares Study 2a and 2b.

The current null results could still be explained by current models of the bilingual lexicon. For instance, the IC model (Green, 1998) would argue that the L1 candidates are being activated but the activation is not strong enough to require inhibition to successfully complete the task. The language task schemes in the IC model (Green, 1998) are moderated by a supervisory attentional system that regulates their activity: by inhibiting the task schema for the non-target language, it can limit interference and maintain activation for the target language. Therefore, in this study where prime presentation is shortened the supervisory attentional system that regulates the activity of the non-target language is not activated as strongly because of the time constraints of the prime. Hence, because it is not being activated as strongly, it is not having a subsequent effect on the task. Whereas in the BIA+ model (Dijkstra & van Heuven, 2002), these null effects are explained by the semantic information from a sentence context that goes through the process of top-down activation from semantic units leading to corresponding orthographic and phonological units. This allows for earlier activation being settled and consequently the process of lexical selection is speeded up. If top-down activation process occurs early enough, it can bypass any influence of co-activated units from the non-target language (Dijkstra & van Heuven, 2002). Therefore, having a sentence constrained and manipulated to one language (e.g., Dijkstra, Van Hell, & Brenders, 2015; Van Hell & De Groot, 2008), as well having the prime shortened, restricts the number of lexical candidates that are activated. Subsequently, the combination of boosted, top-down semantic activation and a restricted set of lexical competitors allow for language selective access of the interlingual homograph L2 representation (Lauro & Schwartz, 2017). However, it could also be that in this experiment we are picking up activation at an earlier stage, and subsequently it is already moving into inhibition. Therefore, the next section runs a comparison between Experiment 2a and 2b to see if the activation of the L1 changes significantly across experiments.

A further aim of the study was to identify whether there were processing differences for in homograph types with a varying degree of orthographic overlap. Overall, all participants were faster at making lexical decisions and made more errors when primed by identical interlingual homographs compared to the near-identical condition, however, there were no significant differences between the groups and therefore no further analysis was conducted.

To conclude, the data suggests that when using an exclusively L2 context and a shortened prime of 200 ms, bilinguals activate the relevant target meaning of the interlingual homograph. Thus, this combination allows for language selective access of the interlingual homograph L2 representation. However, if this is the case, then a future avenue of research would be to use the same paradigm in an exclusively L1 context and a shortened prime duration to see whether the same effects would be present, this would validate the results because it would demonstrate that even in an L1 context when the L1 is more likely to be activated, the bilinguals still do not show any priming towards the non-target language. Thus, it would give support to the idea that it is this combination of biasing the sentence frames to the L2 meaning and using a shortened prime duration that allows the selective access of the target interlingual homograph meaning. The next section tests the difference between the magnitude of the priming durations between Study 2a and 2b.

Study 2(a) and (b) comparison

Study 2a (500 ms prime) and Study 2b (200 ms prime) demonstrated that there was a change from inhibition of the L1 Spanish non-target language to no L1 effects present with the 200 ms prime duration. To explore whether these patterns of results differed significantly from each other, a comparison of the data from the two experiments was carried out.

A mixed ANOVA was conducted where prime duration was entered as a further between-subjects variable. Therefore, the F_1 analyses are a 2 x 2 x (2) x (3) design, with prime durations and group as the between-subject variables, and homograph type, and target as the repeated measures variables. In the F_2 analyses, it is a 2 x (2) x (2) x (3) design, with homograph type acting as the between-subject variable; group, target, and prime duration as the repeated measures variables.

Reaction Time Analysis

The analysis of RTs revealed a main effect of group, F_1 (1, 110) = 37.31, $p <$ .001, η_p2 = .253; F_2 (1, 68) = 466.89, p <.001, η_p2 = .873. Bilinguals took significant longer to respond than the NEs (921 ms vs.726 ms, respectively). There was a main effect of prime duration, F_1 (1, 110) = 12.47, p = .001, η_p2 = .102; F_2 (1, 68) = 466.89, p <.001, η_p2 = .873. Participants were overall faster in Experiment 2a with the 500 ms prime durations compared to Experiment 2b with the 200 ms prime durations (767 ms vs 880 ms, respectively). There was no main effect of homograph type (F_1 (1, 110) = 1.34, p =.250, η_p2 = .012; F_2 (1, 68) = .65, p = .423, η_p2 = .009), but a large main effect of target, F_1 (1.69, 186.38) = 10.01, p <.001, *Greenhouse-Geisser corrected,* η_p2 = .267; F_2 (2, 136) = 10.74, p <.001, η_p2 = .136. Pair-wise comparisons indicated that overall RTs were significantly faster in the English related target condition (783 ms) compared to unrelated targets (832 ms, p <.001, *Bonferroni corrected*) suggesting that overall priming was present. Spanish related targets were significantly slower in comparison to unrelated targets (854 ms, p = .012, *Bonferroni corrected*).

There was a moderate significant interaction between homograph type and prime duration, F_1 (1, 110) = 6.98, p =.009, η_p2 = .060; F_2 (2, 136) = 1.01, p = .366, η_p2 = .015. There was a small significant interaction between target and prime duration, F_1 (2, 220) = 3.50, p =.032, η_p2 = .031; F_2 (2, 136) = 1.01, p = .366, η_p2 = .015. There was also a small significant interaction between target and group, F_1 (2, 220) = 6.73, p =.001, η_p2 = .058; F_2 (2, 136) = 1.01, p = .366, η_p2 = .015. There was a small significant interaction between homograph type and target, F_1 (2, 220) = 3.65, p =.028, η_p2 = .032; F_2 (2, 136) = 1.01, p = .366, η_p2 = .015. There was no significant interaction between homograph type and group.

There was a moderate three-way interaction between homograph type, prime duration, and group, F_1 (1, 110) = 7.73, p =.006, η_p2 = .066; F_2 (2, 136) = 1.01, p = .366, η_p2 = .015. However, there was no three-way interaction found (see Appendix 51). Lastly, there was no four-way interaction between homograph type, target, prime, and group. Since this indicates that the change of prime duration did not significantly affect the patterns of results, no further analyses were carried out.

To summarise, there was an effect of prime duration, with participants being overall faster in the 500 ms prime duration compared to the 200 ms. However, none of the other results indicated that the change of prime duration had significantly affected the ability of the current experimental paradigm to uncover stronger activation/interference from the non-relevant L1.

Error Analysis

Error analysis revealed a moderate main effect of group, F_1 (1, 110) = 11.70, p = .001, η_p2 = .096; F_2 (1, 67) = 12868.22, p <.001, η_p2 = .995. Bilinguals made more errors than NEs (10% vs 6%, respectively). There was no main effect of prime duration F_1 (1, 110) = .54, p < .463, η_p2 = .005; F_2 (1, 67) = 15.89, p <.001, η_p2 = .192. There was a moderate significant main effect of homograph type F_1 (1, 110) = 11.98, p = .001, η_p2 = .096; F_2 (1, 67) = .37, p =.544, η_p2 = .006, with more errors were made on targets that followed identical interlingual homographs prime target condition (9%) compared to those that followed the near-identical ones (9% vs 7%, respectively). There was a large main effect of target, F_1 (2, 220) = 17.80, p < .001, η_p2 = .140; F_2 (2, 134) = 128.62, p <.001, η_p2 = .658. Pair-wise comparisons indicated that fewer errors were made in the English related target condition (6%) compared to the unrelated target condition (9%; p <.001). There were no significant differences between the Spanish related targets (9%) compared to the unrelated targets (p = 1).

There was a moderate significant interaction between homograph type and target, F_1 (2, 220) = 7.49, p = .001, $\eta_{p}2$ = .064; F_2 (2, 134) = .98, p <.379, $\eta_{p}2$ = .014., and a three-way interaction between homograph type, prime duration, and group, F_1 (1, 110) = 8.09, p = .005, $\eta_{p}2$ = .068; F_2 (1, 67) = 4.70, p = .034, $\eta_{p}2$ = .065. None of the other interactions, including the four-way interaction between homograph type, target, prime, and group, were significant (see Appendix 52). Therefore, no further analyses were carried out.

To summarise, there was no effect of prime duration on percentage of errors made. Bilinguals did make more errors compared to NEs, and overall, more errors were made to targets primed by identical interlingual homographs compared to near-identical one. Also, fewer errors were made on the English related targets compared to the unrelated ones, and there was no difference in performance between Spanish related and unrelated targets.

General Discussion

Building of the evidence from Study 1, which had shown activation of both the task-relevant L2 and the non-relevant L1 in a LDT conducted in English, this chapter set out to further explore the interplay between the bilinguals' two languages with a task relying on semantics. The reason for this is that, as we know, interlingual homographs have a special status because they are words identical (or similar) in form across languages, but distinct in terms of meanings. This means that in a LDT, the task could be carried out on the basis of the word form only and therefore semantics may not be required (Grainger & Jacobs, 1996; Playfoot et al., 2018; Poort & Rodd, 2019; Poort et al., 2016). Subsequently, an LDT may be less helpful compared to a semantic task in uncovering the activation of the different languages. Thus, to tap into the meaning(s) of the interlingual homographs, these stimuli were inserted at the end of a sentence frame that provided a semantic context biased towards their L2 readings. Furthermore, the novelty of the current study is that the experiments reported made use of both identical and near-identical interlingual homographs in a systematic way, something that had not been reported in the literature before apart from the study by Di Betta et al. (2015).

In Study 2 (a) it was found that bilinguals behaved differently from NEs, as expected. NEs only activated the English meaning of the interlingual homographs (which are obviously not ambiguous for them), whilst the bilinguals did not show priming for the English related meaning, although this did not reach statistical significance, and significant negative priming for the L1 Spanish meaning. Experiment 2b probed the activation of the two meanings of the interlingual homographs at an earlier point in time, by means of a shortened prime duration. The NEs behaved as predicted and showed priming with the English meaning only. In the case of the bilinguals, a significant priming for the task relevant English meaning was evident, but this time there was no significant activation of the Spanish meaning which may indicate that the bilinguals L1 was not activated. Moreover, the comparison between Experiment 2a and 2b, with the aim of bringing to light significant changes in patterns of results between the two-probe durations, only revealed that overall participants were faster in Experiment 2a compared to Experiment 2b.

Previous research has found that L2 proficiency can influence interlingual homograph processing, in that the more proficient a bilingual is in their L2, the less interference from the non-target language is (e.g., Schwartz & Kroll, 2006; Schulpen, Dijkstra, & Schriefers, 2003; Sunderman & Kroll, 2006). Therefore, in order to examine this possibility further, bilinguals' English proficiency scores were compared between the two experiments. There were no significant differences between LexTALE scores of participants who took part in the two experiments, be they bilinguals (*t* (46) = .23, *p* = .819) or NEs (*t* (46) = -.08, *p* = .934; see Appendix 53). Therefore, differences in proficiency did not contribute to the different findings of the two experiments.

An alternative explanation for the difference in findings, is that 200 ms was not enough time for the bilinguals to access the meaning in the first place, and therefore they did not have to inhibit the Spanish meaning, whereas, in the 500 ms it was long enough to cause L1 interference but enough time for the L1 to be supressed again. A further alternative explanation is that there is a weakened interference effect in the 200 ms prime duration, the priming effects seen in the Spanish related targets compared to the unrelated ones do demonstrate a numerical negative trend, but this is not significant. Thus, it could be that the constraints of the L2 task are having an effect of a reduced negative priming effect, or the manipulation of the prime duration. Therefore, a future research avenue should be to manipulate the prime durations between 200 ms and 500 ms in order to investigate more comprehensively the time-course of the activation of the L1 meaning. Overall, the significant negative priming effect in the 500 ms prime duration and the negative priming trend in the 200 ms would suggest that words are activated in parallel in a non-selective access manner which would support models such as the BIA+ (Dijkstra & van Heuven, 2002).

A further aim of this study was to investigate whether there were processing differences in homograph types with a varying degree of orthographic overlap. For the bilinguals, there were no RTs differences for *homograph type* in the 500 ms prime time condition; though, analysis of the error data showed that bilinguals made more errors in the identical condition compared to the near-identical condition. In the 200 ms prime time condition, bilinguals were faster at making lexical decisions when primed by identical interlingual homographs compared to the near-identical condition, though both showed a trend towards negative priming, and no differences in the error data. More errors in the identical interlingual homographs could be explained by the IC model (Green, 1998) through active inhibition; in the 500 ms prime time-condition the bilinguals see the identical homograph long enough to activate the L1 Spanish meaning but compared to the 200 ms prime-time condition, it gives bilinguals enough time to be able to strongly inhibit the irrelevant L1 meaning; however, this comes at a processing cost leading to more errors. For the near-identical interlingual homographs, the prime being shown is a written word in L2 English only, therefore, it could be that L1 Spanish candidates are being activated but inhibition is not as strong compared to the identical condition which requires more inhibition.

While there were slight differences between prime durations in the behavioural performance of bilinguals when processing interlingual homographs (i.e., more errors in the identical condition compared to near-identical in the 500 ms prime-time condition; shorter RTs in the identical condition compared to near-identical in the 200 ms prime-time condition), the overall pattern of effects was the same for both homograph types when presented with Spanish related targets. In the 500 ms prime-time duration, there was negative priming for Spanish related targets suggesting interference from and subsequent inhibition of the L1 Spanish; and in the 200 ms prime-time duration bilinguals did not show interference from the L1 Spanish related targets as no inhibitory effects were present. These results contrast with findings from the near-identical cognate literature, where the overall magnitude of the cognate facilitation effects is moderated by form overlap, as overlap decreases so does the strength of the facilitative effect (e.g., Dijkstra et al., 2010; Font, 2001; Peeters et al., 2013; van Assche et al., 2011) and thus the facilitation effect is stronger in identical cognates compared to near-identical. Whereas in the present study, even in the near-identical interlingual homographs where the overlap in form was less compared to the identical interlingual homographs, negative priming was still present. In fact, in the 200 ms prime duration near-identical interlingual homographs were 61 ms slower compared to the identical condition for the Spanish-related targets; however, this effect was not significant so should be noted as a trend and not an effect. In the 500 ms prime duration near-identical interlingual homographs were only 22 ms slower compared to the identical condition. Thus, it appears that while the overall magnitude of the inhibitory effects is not moderated by orthographic overlap, they could be moderated by prime duration: with shorter prime durations having a greater inhibitory effect on near-identical interlingual homographs compared to identical.

To conclude, at the shorter 200 ms prime duration, the results would suggest no interference from the L1 language, whereas the longer prime-time duration adds further evidence for co-activation and competition between languages in bilingual language comprehension. The 500 ms prime duration data demonstrate that even in an exclusively L2 context bilinguals activate both L1 and L2 meanings of identical and near-identical interlingual homographs. The results are consistent with the proposal that inhibitory processes are required to resolve the semantic interference and inhibit the irrelevant L1 meaning; but these processes are sensitive to prime duration. The findings indicate that the duration of the prime is an important variable when processing interlingual homographs, with interference being present in the longer 500 ms prime and not the shorter 200 ms. However, it could also be that 200 ms prime duration is not enough time for the bilinguals to access the meaning in the first place, and thus, they did not have inhibit the L1 meaning. In contrast, in the 500 ms prime it was enough time to cause interference but not enough time to supress the L1.

The next chapter moves the focus from ambiguity between languages to ambiguity within a language and makes use of homonyms that are identical in spelling but have distinct meanings. These ambiguous stimuli have striking similarities to interlingual homographs (Poort & Rodd, 2019), in that both have distinct semantics but similar lexical forms. An exploration of ambiguity within a language will provide a comparison with the cross-linguistic ambiguity explored so far and help to paint a more complete picture of what characteristics (e.g., prime duration) play a role in language processing.

Chapter 5 – Barking up the right tree: exploring the effects of homonyms and orthographic neighbours on sentence comprehension.

Introduction

Reading comprehension requires the retrieval of the meaning of individual words to construct a representation of the meaning of a complete sentence (e.g., Davis, 1944; Thorndike, 1973) and relies on the mapping between semantic and orthographic units of representation (Plaut, McClelland, Seidenberg, & Patterson, 1996; Perfetti, 2007). However, the presence of ambiguous words can make the process of reading more complicated, as it adds to the processing demands that are associated with successful language comprehension (Johnsrude & Rodd, 2016). Thus far in this thesis, we have seen that ambiguous words between languages (identical and near-identical interlingual homographs) can have an effect on non-native speakers of English and their L2 processing. The results of Study 1b reported in Chapter 3 indicated that when bilinguals make lexical decisions to interlingual homographs in isolation in their L2 English, linguistic characteristics from both the target L2 and non-target L1 are activated non-selectively for both identical and near-identical Spanish-English interlingual homographs. Chapter 4 extended the exploration by investigating interlingual homographs in a sentence context; and found interference from the L1 non-target language may be influenced by prime duration of the interlingual homograph. At the shorter 200 ms prime duration, the results would suggest no interference from the L1 meaning, whereas in the 500 ms prime duration the data demonstrated that even in an exclusively L2 context, bilinguals activate both L1 and L2 meanings of the identical and near-identical interlingual homographs. The negative priming effect is consistent with the theory that inhibitory processes are required to resolve the semantic interference and inhibit the irrelevant L1 meaning (Green, 1998; Dijkstra & van Heuven, 2002); but these processes are sensitive to prime duration, in that stronger negative priming effects are seen in the later stages of processing compared to the earlier stages (200 ms prime duration).

There are bilingual theoretical frameworks that suggest that, at least above a certain level of proficiency, words from both languages are integrated into a single lexicon (see section 1.2 and Chapter 2 for more details). The findings from Study 1b and Study 2a can be accommodated under an account that assumes a single lexicon. In Chapter 4, it was highlighted that there are some differences between the way Spanish bilinguals process near-identical interlingual homographs compared to Polish bilinguals which may be due to the characteristics of the language or between the stimuli itself (e.g., the Levenshtein Distance was lower in the current thesis compared to Di Betta et al., 2015). Therefore, to overcome those potential issues of varying characteristics between languages, this study focuses on a single language (English; for argument Rogers, Playfoot, & Milton 2018), and finding words that have multiple meanings within a language to attempt to shed some light on what is a general lexical issue, and what is specific to bilingualism, and inhibitory control over languages.

Subsequently, while exploring ambiguity between languages is important for reasons already discussed in this thesis, it is also useful to consider the extent to which monolingual access can aid understanding in the bilingual lexicon. It has been argued that using a monolingual context can help to clarify the predictions that one would make for a bilingual, because there is not as many confounding variables (e.g., another language; Rogers et al., 2018). Furthermore, ambiguity within a language is important to consider as it is commonly encountered in everyday lives; it is estimated that over 80% of English words have multiple meanings (Rodd, Gaskell, & Marslen-Wilson, 2002). Rodd et al. (2002) made an important distinction between two types of ambiguous stimuli within a language; those with multiple related senses, also known as polysemes, such as the word *twist*; and those with multiple unrelated meanings such as the word *bark* (semantically associated with a dog or tree) which will be referred to as homonyms across this research programme. This distinction has implications in terms of processing: Rodd et al. (2002) presented words in isolation in naming tasks and LDTs and found that polysemes were responded to significantly faster compared to homonyms. These results indicate that competition between multiple unrelated meanings of homonyms slows recognition, whereas the processing of polysemes is facilitated due to the rich semantic representation associated with the multiple related meanings.

Poort and Rodd (2019) recently pointed out that there is a large degree of similarity between the pattern of results described for interlingual homographs and homonyms; and that there is a distinct similarity between the stimuli types: both are words with the same spelling but have different meanings (Poort & Rodd, 2019). Generally, interlingual homographs are processed more slowly compared to unambiguous control words, which is commonly referred to as an inhibitory effect, in isolated word paradigms such as LDTs (e.g., Gerard & Scarborough, 1989; De Groot, Delmaar, & Lupker, 2000; Poort, Warren, & Rodd, 2016). Similar patterns of results are also seen in the literature that has used homonyms in isolated paradigms (e.g., Gottlob, Goldinger, Stone, & Van Orden, 1999; Hino & Lupker, 1996; Klepousniotou, 2002; Pexman & Lupker, 1999; Rodd et al., 2002). The combination of the monolingual and bilingual domains' findings suggest that multiple representations of ambiguous stimuli are activated simultaneously; and subsequently, lexical access, at least in isolated word recognition tasks, involves the initial activation of numerous lexical competitors within the lexicon which leads to longer processing times. Thus, there is competition at the semantic level when processing these types of ambiguous words. One potential way to explore these types of stimuli is through a sentence comprehension paradigm, as they provide a context to the reader and allow for the manipulation of the different meanings, therefore making it an effective method to explore these stimuli at the semantic level.

One of the benefits of using a sentence is that it can be deliberately biased towards a particular meaning of a homonym. In the case of an unbiased sentence context, one might expect that the reader would activate both meanings of the ambiguous word as potential candidates. For example, in the example sentence "He located the *bat*" (Vu, Kellas, & Paul, 1998; p. 982), both the words "wooden" and "fly" (Vu et al., 1998; p. 982) might be activated as potential word candidates as the sentence could fit with both meanings of the word *bat* (that is, animal *or* an implement with a handle and a solid surface, commonly made of wood). However, activating both potential word candidates would be costly both in terms of cognitive processing effort and speed, and therefore an inefficient method. When individuals process sentences, they do not wait until the end of the sentence to start accessing word meanings; instead, readers use the information they already have to process sentences as they unfold (e.g., Just & Carpenter, 1980; Marslen-Wilson, 1973; 1975; Traxler & Pickering, 1996). In fact, research suggests that the processing system rapidly selects a single meaning by using the context surrounding the word (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979). However, if the context is not helpful, such as in the sentences that have been left open to both meanings, the language processing system uses a best-guess default solution based on the most common meaning of the word (e.g., Duffy, Morris, & Rayner, 1988; Gadsby, Arnott, & Copland, 2008; Simpson, 1981; Simpson & Burgess, 1985; Twilley & Dixon, 2000) or alternatively on its most recent encounter with the word form (Rodd et al., 2016; Rodd, Cutrin, Kirsch, Millar, & Davis, 2013).

In this experiment (study 3), the aim is to conceptually replicate the experiments in Chapter 4 using a sentence context to investigate how semantic context modulates the activation of the ambiguous stimuli's different meanings, and whether the sentence or manipulation of the prime duration can constrain the meaning that gets activated. Homonyms are used to mimic the role of an identical interlingual homograph, as both stimulus types have distinct semantic meanings, though one within a language (homonym) and the other cross-linguistically (interlingual homographs). In Study 2 of Chapter 4, we saw that manipulating the prime duration to a shorter 200 ms duration, reduced the negative priming in the L1 Spanish related targets, so much so that it was no longer statistically significant. This experiment aims to use the same sentence priming paradigm to compare the pattern of results between experiments; however, whilst homonyms can be used to mimic identical interlingual homographs, there is no clear within-language stimuli parallel with which to compare the near-identical interlingual homographs because native English speakers (NEs) only have one language to access. However, a near-identical condition is essential as without it, it would limit the comparison between the pattern of results of the studies reported in Chapter 4. Subsequently, the study reported in this chapter explores whether utilising orthographic neighbours (ONs) within the English language can mimic the role of a near-identical interlingual homograph. From the ON, a semantically related word was then obtained as shown in the following example: for the homonym *ring*, the most frequent ON is *king*, and its most semantically associated word according to the Latent Semantic Analysis (LSA; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998) is *queen* (see section 5.2 for more information on semantic relatedness; for a visual depiction of the experimental sequence see Figure 11). The study reported in this chapter will therefore use homonyms and ONs as stimuli to see whether the pattern of findings mimics those previously found for identical and near-identical interlingual homographs in Chapter 4.

Homonyms generally have a dominant and non-dominant meaning, and dominance is typically determined by the frequency with which a word meaning occurs (Betts, Gilbert, Cai, Okedara, & Rodd, 2018); for example, the word *vessel* has a high frequency meaning (ship) and a low frequency meaning (a tubular structure carrying blood). Meaning dominance is typically ascertained using a simple word association task where participants are asked to read a word and respond with the first meaning that comes to mind; because presenting an isolated word is absent of semantic context; it is argued that participants are biased to retrieve the word's more frequent meaning (Twilley, Dixon, Taylor, & Clark, 1994). Therefore, when an ambiguous word occurs in a neutral sentence context where both meanings are probable (e.g., the *vessel* in the…), readers appear biased toward retrieving the more dominant meaning (Duffy et al., 1988; Twilley & Dixon, 2000; Rayner & Duffy, 1986). The use of meaning dominance reflects an optimal strategy in word interpretation for the reader. This method of comprehension is advantageous in that it is rapid; however, to successfully understand ambiguous sentences, readers need to be able to detect errors in the meaning of a sentence, and thus trigger appropriate recovery procedures to restore meaning coherence (Blott, Rodd, Ferreira, & Warren, 2020). If a reader is reading too quickly, they may miss crucial disambiguating information. For instance, when a reader reads the sentence "Sally worried the *ball* was going to be too crowded for her liking" (Blott et al., 2020, pp 3), readers will initially activate and integrate the most dominant meaning of the ambiguous word, which is "spherical toy" (Blott et al., 2020, pp 4). However, successful comprehension requires the reader to detect a violation when disambiguating the sentence, and subsequently trigger appropriate revision processes to activate and integrate the intended subordinate meaning of the ambiguous word into the sentence context (in this case a formal dance). Blott et al. (2020) explain that in the example sentence above, there is no prior disambiguating context available, and therefore, the processing system will initially be led down a metaphorical "garden path" (pp 4). The garden-path effect will continue until the reader encounters the disambiguating information (i.e., the word "crowded") in the sentence. At that point, the system will need to detect a violation in the dominant meaning because the selected meaning of '*ball'* is incompatible with the context provided by the word 'crowded'. The reader will then identify that the ambiguity in the word '*ball*' is the reason for the comprehension difficulties, access its previously discarded subordinate meaning, and then integrate the correct associated semantics successfully into the sentence context. The recovery from ambiguous words in garden-path sentences such as in the example given above has been associated with processing costs. For instance, readers generally take longer to process sentences with homonyms as they may require the revision of a meaning selection compared to sentences that contain a matched unambiguous word (e.g., Dopkins, Morris, & Rayner, 1992; Duffy et al., 1988; Rayner & Duffy, 1986; Rayner & Frazier, 1989; Rayner, Pacht, & Duffy, 1994; Sereno, Pacht, & Rayner, 1992). These findings are also supported by eye-tracking studies which have found that readers spend longer fixating within the regions of a sentence that contain disambiguating information, suggesting the detection of a consistency violation (e.g., Ferreira & Clifton 1986; Ferreira & Henderson 1990).

Research has shown that the dominant meaning of the homonym is the default interpretation unless there is a sentence context to steer the interpretation towards the other non-dominant meaning (e.g., Chen & Boland, 2008; Colbert-Getz & Cook, 2013; Foss, 1970; Kotchoubey & El-Khoury, 2014; Lucas, 1999; Martin, Vu, Kellas, & Metcalf, 1999; Rayner & Duffy, 1986; Simpson, 1981; for an overview see Vitello & Rodd, 2015). Thus, context can make the non-dominant meaning as accessible as the dominant one. For instance, Rodd et al. (2013) biased the meanings of homonyms to the non-dominant meaning. Participants first took part in a semantic relatedness task with the aim to expose them to the sentences that contained the homonyms. Participants were instructed to listen to a sentence which was manipulated to the non-dominant meaning of the homonym. After hearing the sentence, a printed probe appeared on the screen, and participants were asked to indicate whether they believed the sentence and probe to be related or unrelated. For instance, an example of a related condition would be the sentence “A *bug* was used to tap the apartment” followed by the probe “secret”; an example of an unrelated condition would be the sentence “A bar was used to smash the pane of glass” and the probe “pencil” (Rodd et al., 2013). After the semantic relatedness task, participants conducted a digit span task as a filler task. The final task was a word association task which was conducted 20 minutes after the first task (i.e., after hearing the sentences), in which the aim was to measure the participants’ preference for the different meanings of the homonyms. The results revealed that participants were 30-40% more likely to interpret the ambiguous word with the non-dominant primed meaning. These findings suggest that even a single encounter with the non-dominant meaning can strengthen the connection between the word-form representation and the primed meaning, such that the non-dominant meaning becomes more readily available compared to the dominant one (Rodd et al., 2013). This shows that the selection of the homonym meaning, is not only influenced by dominance but also by prior exposure of the other non-dominant meaning of the homonym. Therefore, in the current study participants will not be given any prior exposure to the homonym’s meanings. The first-time participants will be exposed to the homonym will be as a prime in the written sentence priming task which will prime the non-dominant meaning of the homonym. The non-dominant meaning has been chosen because it is more akin to the experiments in Study 2 where the sentences were manipulated to bias the L2 meaning of the non-native’s languages.

Chapter 2 discussed the models related to the storage and lexical access debate in bilinguals; however, although the focus is on the bilingual lexicon, it is still relevant to this chapter as how individuals store meanings and access them is not limited to the bilingualism area. Resolving ambiguity depends on access to lexical and/or semantic information, so how this information is stored by individuals may be critical to successfully solve the ambiguity problem both between and within languages. In fact, understanding how ambiguity is resolved in the monolingual domain, could help to understand ambiguity effects between languages by providing a baseline to compare the effects to. As such, the "separate entry" model (also known as selective view; Langacker, 1987) proposes that each meaning is stored separately; this is supported by studies using homonyms in constrained sentences which have found that readers activate the intended meaning only and not the other non-target meaning (e.g., Glucksberg, Kreuz, & Rho, 1986; Oden & Spira, 1983; Simpson, 1981).

It should be noted that it is more common for the non-dominant meaning of a homonym to be successfully inhibited in a constrained sentence compared to the dominant meaning (e.g., Simpson, 1981; Tabossi, 1988; 1989; Tabossi & Zardon, 1993). This is because the dominant meaning is more readily available as it is used more frequently (Langacker, 1987). However, whilst the separate entry model argues that the meanings of homonyms are stored separately, the orthographic and/or phonological forms do overlap, and therefore, this mapping of form can explain why ambiguous words presented in isolation have demonstrated activation of multiple meanings (e.g., Gottlob, Goldinger, Stone, & Van Orden, 1999; Hino & Lupker, 1996; Pexman & Lupker, 1999; Rodd et al., 2002).

This idea of separate semantic entries is generally inconsistent with the bilingual models discussed in Chapter 2 such as the BIA+ model (Dijkstra & van Heuven, 2002) that suggests that bilinguals have an integrated lexicon for the meanings of words. However, bilingual models commonly suggest that the inhibitory effects seen with interlingual homographs in tasks such as LDTs (e.g., Gerard & Scarborough, 1989; De Groot et al., 2000; Poort et al., 2016) may be due to active inhibition of the non-target meaning. The “Separate Entry” model would explain these interference effects as being due to the overlap in orthographic and/or phonological form (Langacker, 1987). Whilst these theoretical accounts can provide a way to understand and predict how homonyms are represented and processed by individuals; a definite answer of how multiple semantic meanings of words are represented, activated, and ultimately selected has still not been found in either the bilingual or the monolingual domain.

The present experiment: Study 3

This study consists of two norming studies detailed in the stimuli section, and one main experiment. The main experiment is methodologically identical to that described in Chapter 4 as it is a written sentence priming task followed by a LDT. However, the experiments differ in stimuli and therefore this makes the main experiment a conceptual replication of the study reported in Chapter 4 (see Figure 1 in Chapter 1 for a visual conceptualisation of the interlingual homograph and homonym stimulus comparison). In both studies, sentences are manipulated to the non-dominant meaning of the ambiguous stimuli (e.g., the non-dominant meaning in Study 2 was the bilingual's L2); and therefore, investigate how people access and manage the meanings of ambiguous words with two distinct meanings, and whether a sentence context can inhibit the dominant meaning.

In Study 2 participants made lexical decisions to targets that were related to the dominant meaning (L1), non-dominant meaning (L2), or unrelated to either meaning. In this study, participants will also make lexical decisions to targets that are either related to one of the two meanings of the homonyms, related ON condition (discussed in more detail later in this section), and unrelated to either meaning. See Figure 11 for a visual illustration of the experimental sequence.

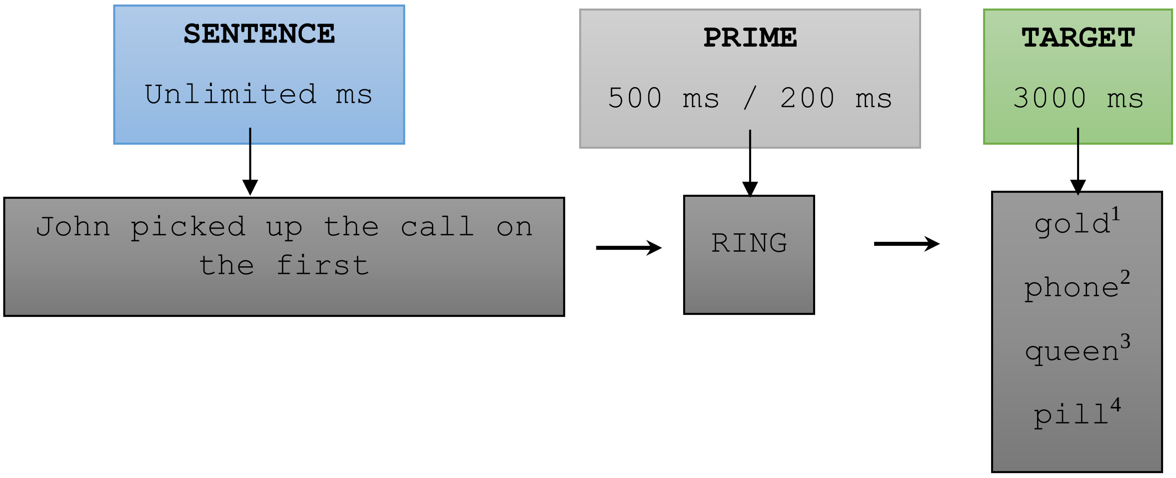

Figure 11.

Sequence flowchart of experiment (Study 3).

In Figure 11, you can see that before each trial begins a 1000 ms blank screen appears followed by a 100 ms fixation cross (+). The prime is the homonym, and targets would either be related to: [1] dominant meaning of the homonym, [2] non-dominant meaning (which the sentence meaning is manipulated to), [3] related-ON condition (for the homonym *ring*, the most frequent ON is *king*, and its most semantically associated word according to the LSA (Landauer & Dumais, 1997; Landauer et al., 1998) is *queen* [4] unrelated to either homonym meanings. There were 4 versions of the experiment, participants would only see one of the target examples in the figure.

There are three aims in this study. Firstly, to investigate how semantic context modulates the activation of the homonyms' different meanings and whether it can constrain the meaning that gets activated. Secondly, to explore whether different patterns of meaning activation can be probed using different prime durations: 200 ms and 500 ms. The 200 ms prime-duration was chosen because it has been argued to be the optimal time duration to capture automatic processing (e.g., Conrad, 1974; Elston-Güttler & Friederici, 2005; Love & Swinney, 1996;

Onifer & Swinney, 1981; Tanenhaus et al., 1979). Instead, the 500 ms prime duration was chosen to ensure that it was long enough to allow for possible meaning activation in participants, but short enough to avoid the use of post-lexical access checking and the use of conscious strategies when making the lexical decisions. Using a prime duration period that is longer than 500 ms has been argued to increase the likelihood that participants will be able to predict the target from the prime (e.g., De Groot, 1984; De Groot & Nas, 1991; Keatley et al., 1994; Neely, Keefe, & Ross, 1989).

Therefore, it is predicted that in the 500 ms prime duration condition NEs will settle on the relevant meaning of the homonym, whereas, in the 200 ms prime duration one may be able to see activation of both meanings. Support for these predictions can be found, for instance, in the study by Elston-Güttler and Friederici (2005) who used sentences to prime the dominant or subordinate meanings of homonyms in native and non-native participants. Using the same paradigm adopted in the current research, homonyms were used as primes and followed by a target for lexical decision. Findings showed that when the prime was onscreen for 200 ms, natives and non-natives participants showed priming for both meanings of the homonym whether contextually appropriate or not. Whereas, when the prime duration was altered to 500 ms, both groups showed priming effects for *only* the contextually appropriate targets in the behavioural data. This study also measured event-related potentials (ERPs), which were generally consistent with the behavioural data. It should be noted, though, that, in the 500 ms condition, non-native participants showed neurophysiological signs of activation of the non-target meaning despite the behavioural data suggesting that only the target meaning was activated.

Thirdly, while homonyms' dominant and non-dominant meanings have been previously investigated in a sentence context, one of the novelties of the current study is that it aims to see whether the homonyms most frequent ONs can be used to mimic an L2 near-identical interlingual homograph. Based on the results of Study 2, where no effects were found in the 200 ms prime condition but inhibitory effects were found in the 500 ms prime condition for interlingual homographs, one may predict that there will be no priming for the target related to the ON compared to the unrelated target in the 200 ms prime condition as it would be too short to activate the neighbourhood. However, in the 500 ms prime condition there will be inhibitory (negative priming) or facilitatory (positive priming) effects. Facilitatory effects might be expected based on Di Betta et al.'s (2015) work which specifically manipulated the orthographic overlap of Polish-English interlingual homographs presented in a L2 sentence context to investigate the performance of proficient Polish-English bilinguals in a sentence priming with LDT. The task was exactly the same as Study 2 and the current study: interlingual homographs acted as primes at the end of the sentence and were followed by targets that were either related to the English, Polish, or unrelated to the meaning of prime. Although the sentence was biased to the L2 meaning of the interlingual homographs, bilinguals demonstrated priming effects with both English and Polish targets, thus suggesting activation of both relevant and non-relevant languages with a 500 ms presentation time of the prime.

The key finding was the differences in how bilinguals processed the two types of interlingual homographs. Identical interlingual homographs demonstrated negative priming, evidence of inter-language competition and inhibition, whereas the near-identical interlingual homographs produced a positive priming effect suggesting that the L1 is less inhibited in this condition. Therefore, in the current study positive priming may also be seen, or alternatively negative priming might be observed in the same manner as Study 2's results.

Before creating the experimental sentences, the dominant and the non-dominant meanings of the selected homonym needed to be established. Therefore, the first norming study outlined in section 5.1 aimed to establish the dominant and non-dominant meaning of a set of previously well-known published English homonyms (Twilley, Dixon, Taylor, & Clark, 1994). The norms by Twilley et al., (1994) were not used as they are outdated and based on the answers of an American sample, whereas this study's participants are from the United Kingdom (UK), and variations between the American and English language are well-documented (e.g., Tottie, 2009; Kutateladze, 2015). Furthermore, at the time of designing this experiment there were no recent norming databases that used a British sample to establish dominance in a set of homonyms. Therefore, in order to select the stimuli and determine dominance for the main experiment, a word association task was conducted to obtain a current representation of the different meanings of a set of homonyms. Since then, it should be noted that there has been a published homonym norming database using British samples by Maciejewski and Klepousniotou (2016), though not all the homonyms used in this study are included in their database. In fact, only six out of the forty of the experimental homonyms are present in that database; though of those six, the dominant and non-dominant meanings are consistent with what is presented in the current study except for one homonym: *lean*. In the current study *lean*'s dominant meaning was "slim with little body fat" whereas in Maciejewski and Klepousniotou (2016) this was their second meaning. While generally the dominance was consistent with this study, it does highlight the importance of conducting a word association survey to ensure that the dominant and non-dominant meaning is reflective of the sample one intends to recruit (see Fitzpatrick, Playfoot, Wray, & Wright, 2015, for recommendation that normative data should be collected per study)

The second norming study outlined in section 5.2 aims to provide a subjective measure of relatedness between the prime and targets, and thus ensure that all pairs of words are appropriately related or unrelated, as needed in the experiment. This second norming study also checks the comprehensibility of the sentences that will be implemented in the main experiment to ensure they can be understood clearly, and that any effects found in the main experiment can be attributed to the stimuli manipulations and not because of poor sentence comprehension.

Method

Participants

A total of 108 monolingual NEs took part in the study, 60 in the 500 ms prime duration condition (13 males and 47 females; *mean age* = 22 years, *SD* = 6.6 years), and 48 in the 200 ms prime duration condition (9 males and 39 females; *mean age* = 21 years, *SD* = 6.3 years). An independent t-test demonstrated that there were no significant differences in age between the groups (*p* = .926).

A total of 13 had no experience with another language (5 in 500 ms; 8 in 200 ms), 79 had experience with one other language (44 in 500 ms; 35 in 200 ms) and 16 had experience with more than one language (11 in 500 ms; 5 in 200 ms). Those participants with experience in other languages stated that they were no longer actively learning a language. Participants completed the LexTALE (Lemhöfer & Broersma, 2012) which is a proficiency questionnaire that is usually used to assess L2 proficiency; however, it is a good assessment of English vocabulary knowledge and has also been used previously to measure proficiency in an individual’s L1 (e.g., Poort & Rodd, 2019). The average score was 92.5% (*SD* = 6.6) for the 500 ms group, and 94.7% (*SD* = 3.5) for the 200 ms group; a t-test showed that there was a significant difference between the proficiency (*t* (106) = 2.34, *p* = .021). However, there was only a difference of 2.2%, therefore, despite the statistically significant difference, both groups' scores indicate that participants were highly proficient in the English language, as one would expect.

Participants were invited to take part in the study via e-mail, in person, through online advertisements and physical posters. Participants on the Psychology undergraduate course at Sheffield Hallam were offered course credits for their participation. Participants had to be native English speakers (NEs), have no known reading-related difficulties (e.g., dyslexia) and not be actively learning another language. The experiments outlined in this chapter received ethical approval from Sheffield Hallam's University Research Ethics Committee (AM/KW/D&S-345).

Stimuli

To prepare the experimental stimuli, two norming studies were conducted and are outlined below in section 5.1 and 5.2.

5.1 Word Association Norming Survey.

The aim of this study was to identify the dominant and non-dominant meaning of a set of 109 homonyms that were obtained from the Twilley, Dixon, Taylor, and Clark, (1994) norms list. Previous word association lists vary across age groups and countries (e.g., Fitzpatrick et al., 2015; Twilley et al., 1994). Therefore, in order to select the stimuli for this study's written sentence priming task, a word association task was conducted to obtain a current representation of the different meanings of a set of previously published English homonyms (Twilley et al., 1994). An example of how meanings can change can be demonstrated when comparing responses from an American sample (Twilley et al., 1994) to the current responses. In the American sample, the popular response to the cue “PORT” was “boat”, the secondary response was “wine”, and the third “bow". None of the participants in this pilot study provided “bow” as a response; the popular response was “wine” and the secondary response was “dock”. The samples differ in time (over two decades apart) and the country in which the survey was conducted in; these factors appear to be enough to influence the dominant meaning of the cue words.

A total of 107 participants were recruited for the word association survey (79 females, 26 males, and 2 preferred not to say; *mean* age = 30.37 years, *SD* = 13.38 years). All participants were monolingual native English speakers (NEs; 104 British nationality and 3 other), were not actively learning another language, and had not been diagnosed with a language impairment (e.g., dyslexia). Despite none of the participants currently actively learning a language, 51 had previous experience with one and 28 had experience with more than one language.

After reading the information sheet and giving consent to take part in the study, the survey asked respondents for the first meaning of the word that came to mind on presentation of a target and allowed for a second or third associated meaning of the word (see Appendix 54). This was either presented on paper or online using Qualtrics (Version January-March 2017). On average, the survey took 40 minutes to complete.

Before analysing the data, participant's responses were cleaned. Firstly, spelling mistakes were corrected only when it was clear that the intended word had been mistyped (e.g., for the cue WAVE, "*ocan*" was corrected to *ocean*). Most of the responses were single words, however, participants occasionally wrote two or more words or a short phrase. Where phrases could be interpreted as logical sequences with a single meaning they were categorised as so (e.g., for the cue TAG, "*attach a label to something*" was categorised as *label*; Wray, 2002). To avoid having many similar labels for the cue words, categorisation was used, and categories were created after looking at all participants' responses to determine the themes. For example, for the cue word PEN, the categories were *"writing"* (example of participants responses: BIC, stationary item, to write with, ink) and *"enclosed space"* (e.g., pig pen, place to keep animals, fenced space, hutch). Moreover, where words had different classifications but similar meanings these were placed into the same category (e.g., for the cue NAIL, *body part* was the popular category meaning and *varnish* was classed as a third/other meaning, but because varnish is commonly used on nails it was placed in the *body part* category; Wray, 2002).

Finally, for the homonyms to be selected for the main experiment they needed to have two distinct meanings. The number of instances of each response for each homonym were counted and turned into a percentage. Therefore, if a third meaning was obtained and it scored higher than 25% then the homonym was not picked for the main experiment. For example, for the homonym GRASS, the dominant meaning was *plant* (90.5% responses), and the second meaning *person (40.4%)*; however, it obtained a third meaning *drugs* (27%). Hence, the homonym GRASS would not have been picked. Furthermore, due to the nature of this chapter's experiment (i.e., making use of an ON) a total of 14 cues were rejected because they did not have any ONs. Subsequently, from the 109 homonyms, a total of 40 were selected for the main experiment, these are outlined in Table 15.

Table 15.

The percentage for the experimental homonym's meanings used in the main experiment.

Item	Homonym	First meaning	%	Second meaning	%	Other meaning(s)	%
1	BAND	music	84	strip	54	wedding bang	11
2	BARK	dog	63	tree	56	*NA*	
3	BAT	animal	71	stick	64	eyelashes	2
4	BLUE	colour	79	depressed	65	calm	2
5	BOIL	heat liquid (cooking)	89	swelling/eruption of the skin	59	be angry	8
6	BOX	container	82	sport	56	square	6
7	CALF	baby cow	65	body part of leg	55	*NA*	
8	CASE	container	67	court	65	*NA*	
9	DOUGH	food	92	money	71	Homer Simpson catchphrase	4
10	DROP	fall	73	small quantity/liquid/water drop	41	take drugs beat drops	4 4

11	DUCK	animal	89	verb to duck/lower head	66	*NA*	
12	FALL	accident	82	autumn	63	*NA*	
13	FAST	speed	88	not eating	59	*NA*	
14	FIRM	something solid/hard	68	business	55	stern/strict	16
15	FOOT	body part	73	measurement	58	bottom/base poetry	8 3
16	LEAN	thin (no fat)	51	to slope	49	*NA*	
17	MIGHT	maybe	67	strength	61	*NA*	
18	MOLE	animal	71	spot	56	person measurement	16 4
19	NAIL	body part	58	metal	62	punch get something right	3 7
20	NOVEL	book/story	80	original	67	*NA*	
21	ODD	strange	68	not even	56	*NA*	
22	PAWN	trade in	55	chess piece	40	porn	7

23	PEN	writing	94	enclosed space	62	*NA*	
24	POACH	cook (e.g., eggs)	62	take/steal associated with animals	59	*NA*	
25	POKER	game	94	tool/fire poking rod	59	Lady Gaga	1
26	PORT	wine	50	dock	56	computer/wire	8
27	PUNCH	hit	79	drink	60	tool	6
28	RANK	position/hierarchy/order	75	extreme (especially to something bad)	52	taxi	3
29	RING	jewellery	57	phone/sound	41	circle boxing	18 3
30	ROCK	stone	76	music	48	person to rely on Dwayne Johnson game	2 1 2
31	SAW	past tense of see	48	DIY tool	65	game	2
32	SPEED	movement	94	drug	68	*NA*	

33	TAG	label	63	game	46	electronic monitoring	5
34	TOAST	bread/heat	95	drink/celebration	65	*NA*	
35	TRAIN	vehicle	75	prepare yourself or someone else for a job/event	55	part of a dress	5
36	TRIM	cut back	84	slim	52	decoration	7
37	VENT	opening (e.g., air hole)	52	express feelings	56	*NA*	
38	WATCH	time piece small clock	56	look at/looking	53	*NA*	
39	WAVE	ocean	51	gesture	52	sound	2
40	WELL	deep hole in the ground	51	healthy	47	done	4

Note – *NA* = not applicable so there were no other meanings reported. The lines signify where homonyms had more than 3 meanings for easy identification for the reader. The homonym meaning averages were 72% for the first meaning, 57% for the second meaning, and 4% for the other meanings.

5.2 Semantic Relatedness and Sentence Comprehension Norms.

The aim of this second norming study was to ensure that the sentences and target pairs used in the main experiment were appropriately related, and that sentences were easy to understand and thus had good comprehension. In Study 2 in Chapter 4 LSA (Landauer & Dumais, 1997) was used as an objective measure to test the relatedness of experimental prime and target pairs as it is a technique in natural language processing that analyses relationships between pairs of words by producing values through an online database (Landauer, Foltz, & Laham, 1998). LSA uses a scale of 0-1, with values closer to 0 indicating the two words are semantically dissimilar, and those closer to 1 indicating that words can be considered semantically similar. However, unlike in chapter 4, this chapter could not utilise LSA because the primes are ambiguous within a language, and subsequently because homonyms have two semantic meanings the LSA is unable to give an accurate semantic score as it affects the strength of the associated meanings. For instance, using one of Twilley et al.'s (1994) homonym and meanings as an example, we can see that for the word *duck* which is associated with an animal and/or an action verb, when comparing it to “quack” and “under” they receive scores of 0.05 and 0.13 on the LSA database which would suggest that these meanings are not semantically similar as they are closer to the value of 0. Furthermore, because homonyms have two meanings, and the sentences were manipulated to the non-dominant meaning of the homonym, we wanted to ensure that participants understood that the sentence and non-dominant targets were related. For example, for the homonym PEN the dominant meaning is an instrument for writing with ink, but the non-dominant meaning is related to a small enclosure in which farm animals are commonly kept. Subsequently for the sentence “On the farm, the pigs were kept in a large PEN” the non-dominant meaning target *cage* would be ideally be rated as related, whereas the dominant meaning target *ink* would be rated as unrelated. The related-ON target *nine* and unrelated target *bread* would also be ideally rated as unrelated to the sentence.

The related-ON target condition was created by calculating the most frequent ON of the homonym, for instance for the word PEN there are 20 ONs (Marian et al., 2012) and the most frequent is *ten*, which is semantically related to the word *nine*, and would be the target word. LSA was used to find a word semantically related to the ON (Landauer & Dumais, 1997).

A total of 80 NEs were recruited from Sheffield Hallam University, with 20 participants in each version of the survey (67 females and 13 males; *mean* age = 21 years, *SD* = 4.1). A total of 48 participants had experience with one and 8 had experience with more than one language. All participants were educated to at least A-level or equivalent. Those participants with knowledge of another language stated that they were no longer actively learning an L2 and had no knowledge of the Spanish language. On average, the survey took 15 minutes to complete.

Four versions of the survey were conducted, as there are four conditions to the study (see Figure 11), participants only saw one version of the survey. The relatedness survey asked participants to rate 40 experimental sentences and target words on being either related or unrelated with each other on a 5-point Likert scale (*1* = strongly unrelated, *2* = unrelated, *3* = neutral, *4* = related, *5* = strongly related; see Appendix 55). NEs were also asked to rate the sentences for comprehension on a Likert scale (*1* = very poor, *2* = poor, *3* = neutral, *4* = good, *5* = very good), and 77% of sentences were rated as very good and 23% as good.

A repeated measures ANOVA demonstrated that there was a significant main effect of relatedness (*F* (1.48, 57.74) = 1289.91, *p* < .001¸*Greenhouse-Geisser corrected*). Greenhouse-Geisser is reported to correct for the violation in sphericity (Dancey & Reidy, 2007). The findings were as expected, participants rated the non-dominant meaning and sentences as being related (*average* = 4.4); and the dominant (*average* = 1.4), related-ON (*average* = 1.2), and unrelated targets (*average* = 1.2) as unrelated (see Appendix 56). Table 16 demonstrates the averages for each sentence (including the prime at the end of the sentence) and the associated target. All participants had the opportunity to identify any words that they did not know the meaning of during the study, none were found to be unfamiliar.

Table 16.

Averages of the relatedness task.

Homonyms	Non-Dominant	Dominant	Related ON	Unrelated
BAND	4.8	1.1	1.3	1.1
BARK	4.2	1.4	1.1	1
BAT	4.2	1.2	1.2	1.1
BLUE	4.7	1.4	1.6	1.2
BOIL	4.6	1.5	1	1.1
BOX	4.5	1.2	1	1.3
CALF	4.6	1.5	1.4	1.1
CASE	4.9	1.3	1.3	1.2
DOUGH	4.7	1.4	1.7	1.1
DROP	4.5	1.5	1.4	1.1
DUCK	4.7	1.5	1.5	1.1
FALL	4.2	1.5	1.1	1.1
FAST	4.7	1.3	1.3	1.1
FIRM	4.4	1.2	1.2	1.1
FOOT	4	1.3	1.1	1.2
LEAN	4.1	1.3	1.2	1.1
MIGHT	4.8	1.6	1.2	1.2
MOLE	4.2	1.4	1.4	1.1
NAIL	4.1	1.3	1.1	1.3

NOVEL	4.4	1.5	1.1	1.1
ODD	4.8	1.1	1.2	1
PAWN	4.4	1.5	1.2	1.1
PEN	4.6	1.4	1.3	1
POACH	3.9	1.5	1.1	1.2
POKER	4.2	1.3	1.2	1.3
PORT	4.4	1.3	1.4	1.2
PUNCH	4.7	1.5	1.2	1.1
RANK	4.9	1.2	1.1	1.5
RING	4.7	1.6	1.1	1
ROCK	4	1.3	1.3	1.2
SAW	4.4	1.2	1.1	1.5
SPEED	4.3	1.4	1.5	1.1
TAG	4.8	1.5	1.2	1.2
TOAST	4.4	1.4	1.1	1.2
TRAIN	4.8	1.4	1.2	1.1
TRIM	4.9	1.4	1.2	1.1
VENT	1.8	1.4	1.2	1.1
WATCH	4.4	1.4	1.5	1.2
WAVE	4.5	1.2	1.2	1.3
WELL	4.3	1	1.3	1.2
Averages	4.4	1.4	1.2	1.2
SD	0.5	0.1	0.2	0.1

Note – Targets words were rated as either being related or unrelated with each other on a 5-point Likert scale (1 = strong unrelated, 2 = unrelated, 3 = neutral, 4 = related, 5 = strongly related).

5.3 Written Sentence Priming Task.

Once the dominant meaning of the homonyms and the semantic relatedness of the primes and targets was established, filler sentences, primes, and targets were created (see Appendix 57).

Targets

Table 17 demonstrates the word frequency, word length, and orthographic and phonological neighbourhood averages for the experimental and filler targets. Experimental targets were matched across experimental conditions according to frequency, length, ONs, and PNs; a one-way ANOVA demonstrated that there were no significant differences between the experimental targets in word frequency (F (3, 156) = 2.56, p = .057), word length (F (3, 156) = 1.29, p = .282), ONs (F (3, 156) = 2.26, p = .083), and PNs (F (3, 156) = 0.32, p = .809; see Appendix 58). Filler targets also did not differ to the experimental stimuli in relation to frequency (t (198) = -0.33, p = .739), length (t (198) = 0.42, p = .677), ONs (t (198) = 0.37, p = .716), and PNs (t (198) = 0.89, p = .377; see Appendix 59). Non-words were matched to the experimental stimuli according to length (t (237) = -1.07, p = .286; Appendix 59). The matching of targets allowed for consistency across all target conditions.

Table 17.

Averages for the experimental and filler targets.

	Frequency	Length	ONs	PNs
Dominant	4.5	4.9	5.9	13.7
Non-Dominant	4.4	5.3	5.4	11.1
Related ON	4.8	5.4	3.0	12.3
Unrelated	4.5	4.9	4.8	11.7
Experimental Combined	4.6	5.1	4.8	12.2
Filler words	4.6	5.0	4.5	10.3
Filler non-words	-	5.3	-	-

Note - The frequency measures were taken from SUBTLEX-UK (van Heuven, Mandera, Keuleers, & Brysbaert, 2014); to standardise the frequency scores, the Zipf scale was used 9Zipf, 1936; 1949) where frequency ranges from 1 (1 – 3 low frequency words) to 6 (4 – 6 high frequency words) or 7 (function words, pronouns, and verb forms like “have”). Length was measured as the number of letters. Orthographic and phonological neighbourhood sizes were taken from 'Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities' (CLEARPOND; Marian et al., 2012)

Sentence Frames

Table 18 shows you an example sentence from each condition, whilst the whole list can be found in Appendix 57. There were a total of 40 experimental sentences that were developed to bias the non-dominant meaning of the homonym. Experimental sentences ended with a homonym which acted as the prime, followed by a lexical decision target. The target was either related to the dominant or non-dominant meaning of the homonym, related to the meaning of a close ON of the homonym, or unrelated to the homonyms meaning (see Figure 11 for a visual depiction of this).

Filler sentences were integrated into the task to enable distraction, so that the homonyms were not so apparent. A total of 120 filler sentences together with filler prime and target words were included in the experiment: 80 of these targets were non-words, 20 targets were related to the prime, and 20 were unrelated to the prime. Relatedness was checked using LSA (Landauer & Dumais, 1997; Landauer et al., 1998), so primes with unrelated targets were closer to the value of 0 such as MUSIC – shirt which had a score of 0.02; whereas primes related to the targets were closer to the value of 1 such as SKY – stars had a score of 0.72.

Table 18.
Example sentences of the experimental and filler targets.

		Target			
Experimental sentence	**PRIME**	Non-Dominant	Dominant	Related-ON*	Unrelated
John picked up the call on the first	RING	Phone	gold	Queen	pill

Note: * the highest frequent ON of the homonym “RING” is *king* so the semantically associated word is queen.

Design

This is a mixed design with target (i.e., Related to Dominant, Non-Dominant, related ON, or unrelated) as the repeated measures variable, and prime duration (2 levels: 200 ms and 500 ms) as a between-subject variable. Presentations of trials were randomised for each participant.

Procedure

After reading the information sheet and giving consent to take part in the study, the participants took part in the computer task (see Appendix 60). Stimulus presentation and recording of RTs and accuracy were obtained with E-prime (Version 2.0; Psychology Software Tools, Pittsburgh, PA). At the beginning of each trial, a blank screen appeared for 1000 ms followed by a fixation stimulus (+) that remained in the middle of screen for 1000 ms. The fixation cross was then removed, and a sentence appeared in the middle of the screen until the participant pressed the *spacebar* to continue. Immediately after this, a prime was presented for 500 ms or 200 ms in uppercase letters followed by the target in lowercase letters. The target remained on the screen until the participant responded, though, if the participant did not respond within 3000ms the target was removed from the screen. Lexical decisions were made by pressing the *x* and the *m* keys on the keyboard. Response hand was controlled such that all participants YES responses were carried out with their dominant hand. Presentation of trials was randomised for each participant, and there were two blocks which allowed participants to have a break in the middle of the experiment. Instructions and targets were presented in black, size 18 Courier New font, on a silver background (identical to Chapter 4). Participants completed a series of 8 practice sentence trials before the start of the main experiment, which were made up of sentences, primes, and targets which did not appear in the main experiment, the practice trials provided participants with feedback of whether the lexical decision was correct or incorrect. Participants were able to complete the practice trial multiple times.

Additionally, participants responded to a total of 12 comprehension questions about the sentences read that required a yes- or no-response by pressing one of two response buttons on a standard keyboard. The questions were included in the experimental task as an attention check every 80 trials, to ensure that participants were indeed reading the sentences. Participants were included in the study if they obtained at least 75% in accuracy on the questions. After the computer task was performed, the participants completed a language background questionnaire, proficiency tests, and were asked to identify any experimental words they were unfamiliar with. Overall, the experiment took approximately 45 minutes to complete.

Results

The results are separated into four sections. The first section discusses the data cleaning process and the descriptive statistics, followed by the inferential statistics. The first two inferential sections look at the prime durations separately: 200 and 500 ms. For these sections repeated measures analyses of variances (ANOVAs) are conducted, RTs and accuracy are the dependant variables, and the target condition is the independent variable (i.e., Dominant meaning, Non-dominant meaning, related to ON, and unrelated targets). The last section compares performance in RTs and accuracy across prime durations in a 2*(4) design, with prime duration as the between participants variable, and target as the repeated measures variable. For all inferential analyses conducted in this study, two ANOVAs are performed. One analysis is performed with *participants* (F_1) and the other with *items* as the random factor (F_2). F_1 involves analysis of the target condition as the repeated measures variable; and prime duration as the between variable (500 ms, 200 ms). In the F_2 analyses, analysis of the target condition is the repeated measures variables; and the prime duration is a repeated measures variable. In the few cases in which the F_2 analyses do not coincide with the F_1 analysis, theoretical interpretation of the statistic will be based on the F_1 analysis but should be treated with caution.

Data cleaning and descriptive statistics

For both RTs and accuracy, one non-word was excluded from the analysis as it was found that it was a low frequency word (i.e., dray), therefore there was a total of 79 non-words. For any analysis involving RTs, only correct responses were considered. The following steps were used to clean the data for both prime durations, the percentages reflect these combined. The incorrect responses (3.8% of the data), sentence reading time below 500 ms (0.2% of the data), and the RTs below 200 ms and exceeding a criterion of +/-2.5 SD from an individual participant's mean (2.9% of the data) were excluded from the analysis. To ensure that the sentences had been understood as intended, any homograph meanings that were unknown to the participant were also excluded from the analysis (0.1% of the data).

Descriptive statistics for RTs and accuracy for both prime durations can be found in Table 19. RTs were faster in the non-dominant condition for both prime durations compared to all other target conditions which were responded to similarly. Accuracy was similar across all target conditions, except in the 500 ms prime duration condition where more errors were made to the unrelated conditions compared to the dominant, non-dominant, and related ON target condition.

Table 19.
Mean RTs and errors for homonyms at 500 ms and 200 ms prime durations for all target conditions.

	Prime Duration	**Dominant Meaning**	**Non-Dominant meaning**	**Related ON**	**Unrelated**
Mean RTs ms (*SD*)	500 ms	703 (156.6)	659 (131.3)	691 (148.1)	701 (159.9)
	200 ms	738 (144.0)	719 (151.1)	746 (149.4)	744 (133.9)
Mean Errors % (*SD*)	500 ms	4% (.07)	3% (.07)	2% (.07)	7% (.11)
	200 ms	4% (.07)	2% (.05)	2% (.05)	4% (.08)

Note - The Standard deviation is reported in brackets.

500 ms Prime Duration: RTs and errors

RTS.

For the RTs, there were outliers in some of the experimental variables as identified by the boxplots (Tukey, 1976), thus z-scores were checked to identify any problematic outliers (using the value of ±3.29). There was one problematic outlier;

therefore, the next highest non-outlier number was used plus one-unit increment in the score as a replacement (Tabachnick, Fidell, & Ullman, 2007). The skewness statistic indicated no problem with skewness as no values were above ±1, and the histograms confirmed this. Subsequently, parametric tests were conducted as these assumptions were met.

The reader should note that in the instances where Mauchly's test of sphericity has been violated, Greenhouse-Geisser is reported to correct for the violation of sphericity (Dancey & Reidy, 2007). The analysis revealed a significant main effect of target, F_1 (2.51, 148.03) = 8.12, p <.001, *Greenhouse-Geisser corrected*, partial $\eta_p 2$ = .121; F_2 (3, 111) = 2.24, p = .087, partial $\eta_p 2$ = .057 (see Appendix 61). Pairwise comparisons indicated that participants were significantly faster in the non-dominant condition compared to the unrelated condition (p = .002, *Bonferroni corrected*) suggesting that priming is present. There were no significant differences between either the target related to the dominant meaning, or the target related to the ON, and unrelated conditions (all p = 1, *Bonferroni corrected*). The non-dominant condition was also significantly faster compared to the dominant condition (p < .001, *Bonferroni corrected*), and the related-ON condition (p = .002, *Bonferroni corrected*). However, the main effect of target should be treated with caution and may not be reliable due to the F_2 not reaching significance, and thus being inconsistent with the F_1 analysis.

Errors.

A repeated measures ANOVA was conducted to compare error scores across the four experimental conditions (see Appendix 62). The analysis revealed a main effect of target F_1 (2.28, 134.46) = 6.84, p = .001, *Greenhouse-Geisser corrected*, partial $\eta_p 2$ = .104; F_2 (2.09, 77.40) = 4.28, p = .016, *Greenhouse-Geisser corrected*, partial $\eta_p 2$ =

.104, suggesting that there were differences in accuracy across conditions. Pairwise comparisons revealed that there was a significant difference between the non-dominant condition compared to the unrelated condition (p = .001, *Bonferroni corrected*) with more errors being made in the latter. There were no significant differences between targets related to dominant meaning, ON, and unrelated ones (p >.05).

200 ms Prime Duration: RTs and errors

RTs.

The z-scores demonstrated there were problematic outliers in the non-dominant (p 35) and dominant condition (using the value of ±3.29), thus, to correct this, the use of the next highest/lowest non-outlier number was used plus/minus one-unit increment in the score as a replacement (Tabachnick, Fidell, & Ullman, 2007). The skewness statistic showed that all target variables were above the value +1 and therefore positively skewed. Histograms confirmed the skewness. To correct for the skewness a lg10 transformation was applied (Field, 2013).

There was a significant main effect of target F_1 (3, 141) = 2.91, p =.037, partial $\eta_p 2$ = .058; F_2 (3, 111) = .76, p = .521, partial $\eta_p 2$ = .020 (see Appendix 63). Post hoc tests revealed that RTs were significantly faster in the non-dominant condition compared to the unrelated condition (p = .012) suggesting priming was present. Though, none of the other comparisons between targets reached statistical significance (p > .05).

Errors.

A repeated measures ANOVA was conducted to compare error scores across the four experimental conditions (see Appendix 64). The analysis showed a non-significant main effect of target, F_1 (2.40, 112.95) = 2.05, p =.124, *Greenhouse-Geisser corrected,* $\eta_p 2$ = .048; F_2 (3, 111) = 2.08, p = .107, $\eta_p 2$ = .053, thus no further analysis was conducted.

Prime Duration Comparison RTs and errors

The prime duration has demonstrated that the pattern of results for the 500 ms prime duration and the 200 ms prime duration are similar. However, to explore whether the pattern of results truly differed significantly from each other, a comparison of the data from the two prime durations was carried out to investigate whether the magnitude of the priming effect difference between the two prime durations.

RTs.

The analysis of prime durations revealed no main effect of prime duration F_1 (1, 106) = 2.97, p =.088, $\eta_p 2$ = .027; 132; F_2 (1, 37) = .74, p =.394, $\eta_p 2$ = .020 (see Figure 12). There was however a moderate main effect of Target, F_1 (3, 318) = 7.62, p <.001, $\eta_p 2$ = .067; F_2 (2.32, 85.87) = 10.95, p <.001, *Greenhouse-Geisser corrected*, $\eta_p 2$ = .228. RTs were significantly faster in the non-dominant condition (*Mean* = 690.96) compared to the unrelated condition (*Mean* = 722.53; p < .001) suggesting overall priming between the groups. There were no significant differences among the other conditions (p = 1). Furthermore, there was a non-significant target and prime group interaction F_1 (3, 318) = .90, p =.443, $\eta_p 2$ = .008; F_2 (3, 111) = 2.06, p =.110, $\eta_p 2$ = .053 (see Appendix 65). Since this indicates that the change of prime duration did not significantly affect the patterns of results, no further analyses were carried out.

Though there were no significant differences between the prime durations, Figure 12 shows that priming was present for both prime durations separately for the non-dominant conditions compared to unrelated targets.

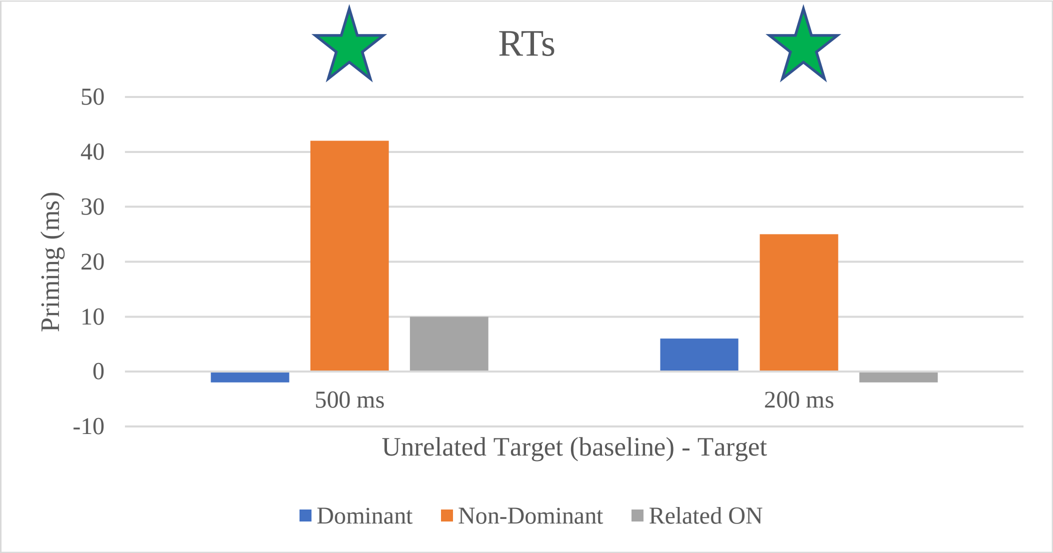

Figure 12.

Priming differences between the baseline's unrelated targets minus the target RTs. Significance is signified with the green stars above the bars.

Errors.

The analysis of prime durations revealed no main effect of prime duration F_1 (1, 106) = 2.06, p =.154, η_p2 = .019; F_2 (1, 37) = 3.40, p =.073, η_p2 =.084. There was however a moderate main effect of Target, F_1 (2.56, 270.90) = 7.27, p <.001, *Greenhouse-Geisser corrected*, η_p2 = .064; F_2 (2.20, 81.27) = 2.05, p =.111, *Greenhouse-Geisser corrected*, η_p2 = .105. Significantly fewer errors were made in the non-dominant condition (*M* = 3% overall errors) compared to the unrelated condition (*M* = 6% overall errors; *p* = .001). Significantly fewer errors were also made in the related ON (9% overall errors) in comparison to the unrelated condition (*p* = .001).

There were no significant differences among the other conditions (all *p's* > .05). Furthermore, there was not a significant target and prime group interaction F_1 (3, 318) = .90, p =.443, $\eta_p 2$ = .008; F_2 (3, 111) = 1.46, p <.230, $\eta_p 2$ = .038 (see Appendix 66). However, the main effect of target should be treated with caution and may not be reliable due to the F_2 not reaching significance, and thus being inconsistent with the F_1 analysis.

Discussion

The broad aim of this study was to explore whether investigating the processing of homonyms in a monolingual domain can shed any light on how identical and near-identical interlingual homographs are processed by bilinguals. This study had three aims. Firstly, to examine how semantic information modulates the activation of the homonyms' different meanings, and specifically whether a sentence context can inhibit the dominant meaning of the homonym. Secondly, to investigate whether manipulating prime duration can highlight different patterns of activation of the homonym's meanings in the target conditions. Lastly, to explore whether an ON can be used to mimic the role of near-identical interlingual homographs and a homonym an identical interlingual homograph.

Homonyms were explored during early (200 ms) and late (500 ms) stages of processing; sentences were biased to the less dominant meaning. It was predicted that at the 500 ms prime duration, participants would activate only the contextually relevant non-dominant meaning of the homonym; whereas in the 200 ms prime duration, it was predicted that priming would be present for both meanings of the homonym, whether contextually appropriate or not. As expected, in the 500 ms prime duration, participants were faster at making lexical decisions and made fewer errors in the non-dominant condition compared to the unrelated condition, suggesting that priming was present for the intended target meaning, suggesting that the sentence context had constrained the selection of the relevant homonym's meaning. Furthermore, in the 200 ms prime duration condition, the same results were obtained for the RTs, though no differences were found in the error data. Overall, these results would suggest that participants were successfully primed by the sentence context towards the non-dominant meaning of the homonym at both early and late stages of processing. Therefore, participants were able successfully to inhibit the irrelevant dominant meaning of the homonym in both prime conditions. This result is consistent with Chapter 4's finding in Study 2a (200 ms prime duration) where bilinguals were successfully primed by the target meaning but did not show interference from the non-target meaning (L1 Spanish). However, it could be that the non-relevant meaning was not activated in the first place in the shorter prime duration, which would mean that inhibition was not required.

This absence of activation of the dominant meaning in the 200 ms prime duration contrasts with a similar study conducted by Elston-Güttler and Friederici (2005). These authors also used homonyms as primes and measured RTs to targets that were preceded by a homonym prime in sentence (e.g., They looked forward to the fun TRIP) or an unrelated prime (e.g., They looked forward to the fun PROGRAM). Targets were either contextually appropriate (e.g., JOURNEY) or inappropriate in meaning (e.g., STUMBLE). In the 200 ms prime duration, Elston-Güttler and Friederici (2005) found that both meanings of the homonym were activated which is line with previous studies who have also obtained the finding of automatic retrieval of both homonym meanings using prime durations of 200 ms or shorter (e.g., Conrad, 1974; Love, Maas, & Swinney, 2003; Onifer & Swinney, 1981; Seidenberg et al., 1982). These findings support the idea of multiple access at an early stage of processing; however, the present findings conflict with these previous ones as only the contextually appropriate meaning was activated.

However, the difference between the current findings and Elston-Güttler and Friederici’s (2005) findings could be attributed to the stimuli used. The authors did note that some of the homonyms included in their stimuli set had more than two distinct meanings, whereas the stimuli for the current study were developed in such a way as to ensure that only homonyms with two clear distinct meanings were used (as validated by the word association survey outlined in section 5.1). Thus, Elston-Güttler and Friederici (2005) may have obtained priming for the dominant condition in the 200 ms prime-duration condition because the homonyms with multiple meanings have a richer semantic network compared to just two meanings. Furthermore, previous literature has found activation of both the dominant and non-dominant meaning of the homonym (e.g., Conrad, 1974; Love et al., 2003; Onifer & Swinney, 1981; Seidenberg et al., 1982) which is in contrast to the current study’s findings. However, as noted Rodd et al. (2002) it could be that this difference is because of the stimuli itself. Rodd et al. (2002) noted that there are two types of ambiguous words, ones with distinct meanings (homonyms) and the other type with multiple related meanings (polysemes). Rodd et al. (2002) presented words in isolation in naming tasks and LDTs and found that polysemes were responded to significantly faster compared to homonyms; this would indicate that competition between multiple unrelated meanings of homonyms slows recognition, whereas polysemes were facilitated due to the rich semantic representation associated with the multiple related meanings. Therefore, while the current study used only homonyms, the difference in results may be due to the stimuli itself.

In addition, the current study differed from previous research investigating homonyms because it introduced a condition that attempted to mimic a second language competitor (i.e., a near-identical interlingual homograph). This was to allow for the cross-comparison of findings between this study's findings and those reported in Chapter 4. Homonyms were used to mimic the role of an identical interlingual homograph, as both stimulus types have distinct meanings; and a homonym’s ON was used to mimic the role of a near-identical interlingual homograph. Following from chapter 4’s results, where no effects were found in the 200 ms prime condition but inhibitory effects were found in the 500 ms prime condition, it was predicted that there would be no effects in the 200 ms prime condition as it would be too fast to activate the neighbourhood. Whereas, in the 500 ms prime condition either an inhibitory or facilitatory effect might be observed. The findings suggest that in both conditions, participants did not perform any differently when they made lexical decisions to the related-ON condition target compared to the unrelated condition. In the current experiment, not finding an effect in the 500 ms prime duration contrasts with the pattern of results seen in Chapter 4 where the non-target dominant meaning was activated (as evidenced by negative priming). Thus, while Poort and Rodd (2019) recently pointed out that there is a large degree of similarity between the pattern of results described for interlingual homographs and homonyms in the literature. The results from this experiment would provide evidence that related-ONs are not able to mimic the pattern of results of near-identical interlingual homographs seen in Study 2a. While there were no related-ON effects in this task, future research could explore this proposition further using different paradigms as it has been noted that ambiguity effects can be influenced by task demands (see Poort & Rodd, 2019, for brief comparison), and so it could be that another task would be able to mimic the near-identical interlingual homograph effects or that related-ON's are not able to mimic near-identical interlingual homographs regardless of the task.

Overall, the findings in this study support the selective access view (McClelland, 1987; Simpson, 1994), that individuals have control over which meanings are activated; participants exhibited control as only the contextually relevant meaning of the homonym was accessed. This finding is supported by previous research that has also found that a homonyms' meaning can be inhibited in a sentence context (e.g., Glucksberg et al., 1986; Oden & Spira, 1983; Simpson, 1981). The selective access view is commonly challenged by studies that have found that only the non-dominant meaning of a homonym can be inhibited by a sentence context (e.g., Simpson, 1981; Tabossi, 1988; 1989); but in Study 3, a sentence context was enough to stop the dominant meaning from interfering with the task. Although why participants were able to inhibit the dominant meaning in the 200 ms prime duration in this chapter is unclear. This prime duration was chosen because it has been argued to be the optimal time duration to capture automatic processing (e.g., Conrad, 1974; Elston-Güttler & Friederici, 2005; Love & Swinney, 1996; Lucas, 1999; Onifer & Swinney, 1981; Tanenhaus et al., 1979). Although, these studies generally focused on participants responding to the homonyms, and not semantically related targets such as in this experiment, so it may be simply due to different task demands.

A future research avenue could be to shorten the prime duration even further to 150 ms, to explore whether this would result in both the dominant and non-dominant meaning of the homonyms being activated. This prime duration has been used in priming research utilising regular words and non-words in a LDT, and has shown a greater facilitatory effect in comparison to prime durations of 700 ms (e.g., Hill, Strube, Roesch-Ely, & Weisbrod, 2002; Kiefer, Weisbrod, Kern, Maier, & Spitzer, 1998). For instance, Hill et al. (2002) asked native German speakers to take part in a LDT where there were 3 different prime-target semantic relationship conditions, and one non-word target condition. Prime-targets were either non-related pairs (e.g., Blatt–Auto, English translation “Leaf-Care”), indirectly related word pairs of a strong semantic associate (e.g., Zitrone–süß, English translation “Lemon–Sweet”), directly semantically related pairs (e.g., Henne–Ei, English translation “Hen-Egg”), or the target was a non-word in German (e.g., Bild–Gerba, English translation “Picture-Gerba”). Most of the indirectly related targets were antonyms of a strong semantic associate (e.g., lemon–[sour]–sweet; square brackets refer to the strong semantic associate) or a property of a strong semantic associate (e.g., lion–[tiger]–stripes) of the prime word. The results revealed that participants were faster in the direct target condition, compared to the indirect, unrelated, and non-word target condition. Moreover, participants were faster in the 150 ms prime duration compared to the 700 ms prime duration in both the direct and indirect semantically related targets. Hill et al. (2002) attributed this effect to the spreading of semantic activation, and this process being the reason that participants were faster in the 150 ms prime duration compared to the 700 ms, whereas, in the later prime duration participants had enough time to select the correct meaning. Therefore, if future research shortened the prime duration from the current experiment’s 200 ms to 150 ms, it may allow for the spreading of semantic activation which will allow for the capturing of both the dominant and non-dominant meaning of the homonym.

To conclude, the findings add support for the selective view (Langacker, 1987) that proposes that each meaning of the homonym is stored separately. The findings from this study have shown that the homonym’s dominant meaning can be inhibited in a constrained sentence that is biased towards the non-dominant meaning at both early (200 ms) and later (500 ms) stages of processing. It has also shown identical interlingual homographs mimic the pattern of results of homonyms in the 200 ms prime duration with both patterns of results not showing interference from the dominant meaning of the ambiguous stimuli. However, in the later processing stages, this study did not mimic the pattern of the results in Study 2a which used a 500 ms prime duration where interference was shown for the dominant L1 meaning. Furthermore, the results from this experiment provide tentative novel evidence that related-ONs are not able to mimic the pattern of results of near-identical interlingual homographs seen in Study 2a. The next and last chapter of this thesis will move the focus to concluding remarks and discuss future avenues of research.

Chapter 6 – Concluding remarks and future work.

6.1 Concluding Remarks

The aim of this thesis was to investigate cross-linguistic ambiguity and the relative activation of the bilinguals' two languages during written word recognition. It focused on assessing the activation of the different meanings of Spanish – English interlingual homographs in both lexical and semantic tasks. Specifically, the goal was to see how activation of the bilinguals' L1 Spanish non-target meaning was modulated by orthographic overlap in an exclusively L2 task. One of the novelties of this research programme is that it makes use of identical and near-identical interlingual homographs which have not been investigated separately in previous research, apart from limited research in semantic tasks (Di Betta, Okurowska, & Morgan, 2015).

Although the focus was on interlingual homographs, this thesis also later explored ambiguity within a language, and therefore investigated the activation of both homonyms meanings in a monolingual semantic task. Homonyms were selected based on the suggestion that these stimuli types can be compared to interlingual homographs (Poort & Rodd, 2019), as both are orthographically identical, but have distinct semantic meanings. Moreover, a further novelty of this research programme is that original stimuli were developed to mirror near-identical interlingual homographs by using orthographic neighbours (ONs) within a language. Each experimental chapter will now be discussed accordingly and outline its contribution to the research programme as a whole.

Chapter 3 introduced a set of Spanish – English identical and near-identical interlingual homographs and collected their psycholinguistic characteristics in both languages in order to be used in conjunction with a lexical task (Study 1). To complement the information that was already available about these stimuli, a study was conducted to collect imageability and AoA scores. This was necessary because, whilst many ratings already exist for these two variables in many languages, fewer are available for *both* meanings of the interlingual homographs; and gathering values from separate sources in each language is problematic methodologically. Therefore, to overcome this problem, we collected these ratings for both English and Spanish to add to our database. This information was added to other information available in the literature to compile a database that includes the stimuli's linguistic properties in both English and Spanish (Study 1a). Therefore, this is a controlled database that can be used in future research.

The last part of chapter 3, Study 1b, examined the impact of these psycholinguistic characteristics on the processing of interlingual homographs when bilinguals make lexical decisions. This task provided insight into the interplay between language activation, by identifying the L1 Spanish and L2 English psycholinguistic variables that best account for the bilinguals' performance when making lexical decisions in an L2 context. The findings showed significant influences of Spanish orthographic neighbourhood on RT, of Spanish frequency, AoA and imageability on accuracy, and of Spanish phonological neighbourhood on both RT and accuracy even though the task was conducted entirely in their L2. Also, of interest is the finding that more errors were made to identical interlingual homographs compared to near-identical; although, this finding was interpreted with the suggestion that these words were identified as Spanish words by the bilinguals and thus rejected as being an English word. What is interesting is that these results should be placed in the context of findings available in the literature that have not been able to highlight L1 effects in a LDT carried out in the participants' L2 (e.g., Dijkstra, De Bruijn, Schriefers, & Ten Brinke, 2000a, Experiment 1; Dijkstra, Timmermans, & Schriefers, 2000b). Instead, Study 1b established that it is possible to bring to light L1 influences without the element of priming. Overall, it was concluded that this experiment provided evidence that there is some activation of the bilinguals' L1 even when performing a task in their L2, and that activation of the bilinguals' L1 Spanish non-target meaning is not modulated by orthographic overlap as there were no significant differences between the homograph types.

A by-product of Chapter 3 is that there will be an easy-accessible list of both identical and near-identical Spanish – English interlingual homographs with varying degrees of spelling overlap, with their associated linguistic characteristics in both languages. This database will enable researchers to select and match stimuli more rigorously in future experiments. In order to appropriately design studies using interlingual homographs as stimuli, being able to manipulate or control these linguistic properties is vital. To ensure that this database can be accessed, a manuscript including the database and LDT experiment outlined in Chapter 3 has been prepared and is ready for submission to the Behavior Research Methods journal.

The findings reported in Chapter 3 made an important contribution to our understanding of interlingual homographs from a linguistic view by providing evidence that L1 linguistic properties are active during the lexical decision process in RTs and accuracy. Chapter 4 extends this investigation of interlingual homographs from an isolated paradigm to a sentence one; interlingual homographs were inserted at the end of a sentence frame that was semantically biased to the L2 readings. The experiments (Study 2) made use of both identical and near-identical interlingual homographs in a systematic way in a sentence context, something that has not been reported in the literature before using the Spanish - English language; it has only been reported in Polish – English bilinguals (Di Betta et al., 2015). An additional novelty is that, instead of using the interlingual homograph as the target, they were used as the prime, and bilinguals made lexical decisions to targets semantically related to the interlingual homograph (i.e., Spanish related, English related, and unrelated). This was done to focus on the meaning(s) of the interlingual homograph, because if they were used as the targets (instead of primes), participants may have made lexical decisions based on the orthography, and thus any interference from the L1 non-target meaning may have been due to orthographic overlap and not because of the distinct semantics meanings.

In Study 2a significant negative priming was found in the bilinguals for the Spanish-related targets and interpreted this as evidence of inhibition of the non-relevant L1 reading of the interlingual homograph. What is interesting, is that in Study 1b the participants made lexical decisions to the interlingual homographs themselves and found evidence of L1 linguistic activation, whereas Study 2a extends this finding of L1 activation because participants made lexical decisions to targets semantically related to the different readings of the interlingual homographs. Therefore, while some researchers argue that a lexical decision can be made on the orthographic form of a word and not the semantics (e.g., Grainger & Jacobs, 1996; Playfoot et al., 2018), Study 2a provides evidence that the role of semantics is active, and this activation is not just limited to the L2, but also to the L1.

The findings from Study 1b compliment the finding from Study 2a, both studies demonstrated that the L1 non-target language is activated in exclusively L2 tasks, which supports the notion discussed in Chapter 2 that languages are activated non-selectively.

The BIA+ model (Dijkstra & van Heuven, 2002) and the Inhibitory Control model (Green, 1998; 2003) can help explain why there was L1 activation in an L2 task (the Multilink Model, Dijkstra et al., 2019, is yet to provide an explanation of how interlingual homographs are structured in the lexicon). In short, both the BIA+ model and the Inhibitory Control model rely on the concept of inhibition, that is, when bilinguals encounter an interlingual homograph, both meanings become activated whether relevant to the task or not. Therefore, in order to successfully complete the task in the intended language, in this case the L2, the L1 non-target meaning must be inhibited. This process of inhibition results in the slower processing times and negative priming seen in Study 2a, and the activation of both meanings of the interlingual homograph can help explain why linguistic properties from both the L1 and L2 were active during lexical decision making. Furthermore, although the Hierarchical Model (Potter et al., 1984) and the RHM (Kroll & Stewart, 1994) opt for a separate lexicon, the attributed L1 interference effects seen for the Spanish related targets can still be explained by these models through the shared language attributes at the conceptual level.

Study 2a primed participants for 500 ms; however, there is evidence from monolingual (Gernsbacher, Varner, & Faust, 1990; Elston-Güttler & Friederici, 2005) and bilingual studies (Beauvillain & Grainger, 1987; Martín, Macizo, & Bajo, 2010) that show that it is possible to reveal the activation of an ambiguous word's different meanings by probing them at different points in time. Therefore, the goal of Study 2b was to use the exact paradigm as Study 2a for comparison ease, but to shorten the prime duration to 200 ms in the hope that this would help capture activation of the L1 meaning of the interlingual homographs more fully and also establish whether the negative priming to the Spanish-related targets was due to inhibition or interference of the L1. In contrast to the 500 ms inhibitory finding, no significant priming was found with the Spanish-related targets. There were also no significant differences between the identical and near-identical interlingual homographs, with both types eliciting the same pattern of results (i.e., negative priming). The model that bests accounts for the absence of L1 activation in Study 2b is the Revised Hierarchical Model (RHM; Kroll & Stewart, 1994) discussed in Chapter 2. The RHM assumes that bilinguals have separate lexicons for each language, and hence no effects are present as they do not interfere with one another. While this explanation may first appear to contradict the findings in Study 2a, the RHM is also able to explain why there are semantic interference effects of the L1 non-target language through the underlying sharing of features of the interlingual homographs at the conceptual level. Though the meanings of interlingual homographs differ across languages, all meanings are stored at the conceptual level, so the connection between the two lexicons through the conceptual level (Luo, Craik, Moreno, & Bialystok, 2013) can lead to interference effects when processing interlingual homographs. Moreover, according to the IC model (Green, 1998) the presence of no significant priming effects to the L1 Spanish-related target would suggest that the bilingual enforced some control over their non-native language. However, the findings from Study 2 may tentatively suggest that interference may be sensitive to prime duration, and that both identical and near-identical interlingual homographs are processed similarly to one another.

A bilingual's continuous management of their two languages competing for attention requires the use of executive function, more often compared to monolinguals who do not have competing languages; thus, a bilinguals' ability to manage their languages enhances functions of executive control (e.g., Bialystok, Craik, & Luk, 2008). It is also suggested that this advantage extends to help slow the natural declining of cognitive control associated with the aging process (Bialystok, Craik, & Klein, 2004; Bialystok, Craik, Klein, & Viswanathan, 2004). For instance, white matter integrity in the brain generally decreases with age resulting in a decline of cognitive function and control. However, bilingualism is associated with higher levels of white matter integrity and connectivity and subsequently greater cognitive ability compared to monolinguals (Luk, Bialystok, Craik, & Grady, 2011). Overall, while in this thesis, bilinguals were slower at reacting compared to monolinguals in these lexical tasks (Study 2a and 2b), the results provide evidence that the bilinguals' L1 was activated in an L2 task. Therefore, supporting the notion of the continual management and maintenance of two competing languages, which may in turn may strengthen pathways in the brain, resulting in a widespread network of white matter connectivity, and help protect against natural cognitive decline (Luk et al., 2011).

So far, this thesis had focused on how non-native speakers of English process interlingual homographs, and while this is vital to understanding of language processing, we wanted to explore ambiguity as a whole, in order to paint a more complete picture of what characteristics play a role in language processing. Therefore, based on the suggestion that homonyms have striking similarities to interlingual homographs (Poort & Rodd, 2019), in Chapter 5, an experiment was developed that was methodologically identical to Study 2 to allow for cross-comparison of the result patterns, and explore homonyms in a sentence context biased to the non-dominant meaning in early and late stages of processing (Study 3). Although, to do this fully stimuli were developed from the orthographic neighbourhood of the homonyms to mimic near-identical interlingual homographs (see Figure 1). Overall, the findings demonstrated that there was significant priming for the non-dominant meaning which showed that participants were successfully primed by the sentence. However, there was no significant evidence of priming of the non-target dominant meaning or the related-ON condition in both early (200 ms) and later (500 ms) stages of processing.

The experiments outlined in Chapter 4 and 5 will be drafted into manuscripts and form the basis of two submissions to peer-reviewed journals such as the journal of Bilingualism: Language and Cognition.

6.2 Future Work

The findings reported in this thesis do present an insightful picture of how identical and near-identical interlingual homographs are processed in both lexical and semantic tasks, in addition to providing an exploration into using ambiguous stimuli within a language to mimic ambiguous stimuli between languages. However, there are several gaps that need to be filled. Firstly, we have demonstrated that the non-target L1 Spanish is activated during L2 lexical decision suggesting non-selective access over bilinguals' languages. A future research avenue to confirm this could be to vary the task instructions to allow a "yes" response for stimuli that are words in either language, which ought to result in an increase in activation of the L1 linguistic properties and an accuracy advantage for identical interlingual homographs compared to near-identical (which would be consistent with our findings). Furthermore, this study could be replicated in a Spanish context, and use proficient bilinguals whose L2 is Spanish, and L1 English, so essentially swapping the L1 and L2. We would expect, based on our findings, that the L1 English characteristics would be activated, this finding would strengthen the findings found in Study 1b.

Secondly, in the written sentence priming tasks using interlingual homographs in Chapter 4, the pattern of results differed depending on whether the language processing was probed at later (500 ms, Study 2a) or earlier activation (200 ms, Study 2b). In Study 2a there was interference from the non-target meaning as evidenced by the significant negative priming, but when the 200 ms prime duration was used, there was no significant interference present from the non-target meaning which suggests that bilinguals were able to activate the relevant target meaning of the interlingual homograph. It remains unclear whether shortening the prime duration allowed for bilinguals to activate only the relevant target meaning. Therefore, a future avenue of research would be to use the same paradigm with a shortened prime duration to see whether the same effects would be present, this would strengthen the notion that shortening the prime duration allows for selective access of the target interlingual homograph meaning.

Another gap related to prime duration is that of our written sentence priming task using homonyms in Chapter 5 where we obtained the same pattern of results in early and late stages of processing. The findings suggested that there was no evidence that the dominant meaning of the homonym or the related-ON was activated. However, it could be the prime was too long to capture both meanings of the homonym, and the related-ON meaning. A future research avenue could be to shorten the prime duration even further to 150 ms to see whether the meanings would become active. This prime duration has been used in priming research that uses regular words and non-words in LDTs and has shown greater facilitatory effects for this prime duration compared to later prime durations (e.g., Hill, Strube, Roesch-Ely, & Weisbrod, 2002; Kiefer, Wisbrod, Ken, Maier, & Spitzer, 1998). This effect can be explained by the spreading of semantic activation (Hill et al., 2002), and therefore, if future research shortened the prime duration from the current experiment's 200 ms to 150 ms, it may allow for the spreading of semantic activation which will allow for the capturing of both the dominant and non-dominant meaning of the homonym.

A further point to note is that the bilinguals who took part in the experiments were all proficient bilinguals, and this was to ensure that the exploration of ambiguity was focused on the stimuli and not the proficiency of participants. Previous research has shown that proficiency can affect language processing, and that interference of the dominant L1 is stronger in less proficient bilinguals when processing interlingual homographs in tasks such as LDTs (e.g., Schulpen, Dijkstra, & Schriefers, 2003; Brenders, Van hell, Dijkstra, 2011). One of the possibilities of this is that during early stages of L2 learning adult learners rely on the L1 and the lexical form of words, and as proficiency increases the reliance of word form shifts to a reliance on meaning (Talamas, Kroll, & Dufour, 1999). Brenders et al., (2011) argue that this is because less proficient bilinguals are confused by the orthographic ambiguity, whereas more proficient bilinguals' resort to different strategies in order to resolve the ambiguity problems such as the semantic readings of the ambiguous word. Therefore, a future research avenue is to recruit less proficient bilinguals and see how they behave in the written sentence priming task (i.e., study 2) which is based on making lexical decisions to semantically related targets, and not the interlingual homographs themselves. If the interlingual homographs were the targets, one would expect there to be an increase in processing times because as mentioned above, the less proficient bilinguals would be relying on the lexical form, and this would therefore be confusing for them. However, less proficient bilinguals would be making lexical decisions to semantically related targets, and therefore the issue lexical overlap is removed, because bilinguals are instead making lexical decisions to semantically related targets. Subsequently, one might expect there to be less interference of the dominant language because less proficient bilinguals are argued to focus more on the orthography compared to the semantics. Conducting this further research, would also validate the findings in our Study 2, and tentatively suggest that bilinguals accessed the semantic reading of the interlingual homograph in the later stages of processing but not the earlier ones.

Lastly, the inconsistency between the results in Study 2a and 2b in how identical and near-identical interlingual homographs are processed compared to those of Di Betta et al. (2015) calls for a greater variety of L1s to be explored in order to paint a more complete picture of all the characteristics at play in bilingual language processing. In Study 2, both interlingual homograph types demonstrated negative priming patterns for the Spanish-related targets compared to unrelated (this was significant in Study 2a), whereas, in Di Betta et al.'s (2015) research facilitatory effects were found for near-identical interlingual homographs, and inhibitory effects for identical. These differences as discussed in Chapter 4, could be because of the underlying differences between the languages, but to test this idea, further L1s need to be investigated. This future research would allow us to further explore which language characteristics play a role in interlingual homograph processing, and how this impacts orthographic overlap.

In conclusion, the findings in this thesis offer a novel and broader perspective with which to investigate interlingual homographs. This thesis provides the most comprehensive assessment of the impact of orthographic overlap on Spanish – English interlingual homograph processing to date. It has focused on the impact of L1 and L2 psycholinguistics variables on interlingual homograph processing; in addition, to investigating interlingual homographs in a L2 context in early and late stages of processing. The thesis also looked at the monolingual domain to explore whether ambiguous stimuli within a language could help us understand more fully how ambiguous stimuli are processed overall. Furthermore, the significance of this research is that it provides a more comprehensive list compared to those available in the literature of identical and near-identical Spanish – English interlingual homographs with their associated psycholinguistic properties for researchers to utilise for their own investigations. It is hoped that this research, with its novel contribution, will stimulate new investigations within linguistic ambiguity and how it is resolved, be that in bilingual or monolingual populations.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242-275.

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, *23*(4), 557-582.

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520-529.

Alario, F. X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140-155.

Alija Fernández, M., & Cuetos Vega, F. (2006). Efectos de las variables léxico-semánticas en el reconocimiento visual de palabras. *Psicothema, 18* (3), 485-491.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 802-814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy*? Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 18(2), 234–254. https://doi.org/10.1037/0278-7393.18.2.234

Assink, E. M., van Well, S., & Knuijt, P. P. N. A. (2003). Contrasting effects of age of acquisition in lexical decision and letter detection. *American Journal of Psychology*, *116*(3), 367-388.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (CD-ROM). Linguistic data consortium. *Philadelphia, PA: University of Pennsylvania*.

Baker, C. (2007). Becoming bilingual through bilingual education. In P. Auer & L. Wei (Eds), *Handbook of multilingualism and multilingual communication* (pp. 131-152).

Balota, D. A. (1994). Visual word recognition. In M. A. Gernsbacher (Ed), *Handbook of psycholinguistics* (pp. 303-358).

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283-316.

Bearden, W. O., Netemeyer, R. G., & Mobley, M. F. (1993). *Multi-item measures for marketing and consumer behaviour research.* Handbook of Marketing Scales.

Beardsmore, H. B. (1986). *Bilingualism: basic principles* (Vol. 1). Multilingual matters.

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of memory and language*, *26*(6), 658-672.

Betts, H. N., Gilbert, R. A., Cai, Z. G., Okedara, Z. B., & Rodd, J. M. (2018). Retuning of lexical-semantic representations: Repetition and spacing effects in word-meaning priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(7), 1130-1150.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, *19*(2), 290-303.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, *19*(2), 290-303.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 859-873.

Birdsong, D. (2016). Dominance in bilingualism: Foundations of measurement, with insights from the study of handedness. In C. Silva-Corvalán & J. Treffers-Daller (Eds), *Language dominance in bilinguals: Issues of measurement and operationalization*, *(*pp. 85-105).

Birdsong, D., & Vanhove, J. (2016). Age of second language acquisition: Critical periods and social concerns. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 163–181). American Psychological Association. https://doi.org/10.1037/14939-010

Bleasdale, F. A. (1987). Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(4), 582-594.

Bloomfield, L. 1933. Language. New York: Holt, Rinehart & Winston.

Blott, L. M., Rodd, J. M., Ferreira, F., & Warren, J. E. (2020). Recovery from misinterpretations during online sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0000936

Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in Stimulus–Stimulus inhibition. *Bilingualism: Language and Cognition, 17*(3), 610-629.

Bonin, P., Chalard, M., Méot, A., & Barry, C. (2006). Are age-of-acquisition effects on object naming due simply to differences in object recognition? Comments on Levelt (2002). *Memory & Cognition, 34*(5), 1172-1182.

Bowers, J. S., Mimouni, Z., & Arguin, M. (2000). Orthography plays a critical role in cognate priming: Evidence from French/English and Arabic/French cognates. *Memory & Cognition*, *28*(8), 1289-1296.

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences, 16*(2), 106-113.

Breedin, S. D., Saffran, E. M., & Coslett, H. B. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology, 11*(6), 617-660.

Brenders, P., Van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology*, *109*(4), 383-396.

Briederman, G. (1966). The recognition of tachistoscopically presented five-letter words as a function of digram frecuency. *Journal of Verbal Learning and Verbal Behavior*, *5*, 208-209.

Broadbent, D. E., & Gregory, M. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning & Verbal Behavior*.

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms?. *Quarterly Journal of Experimental Psychology*, *64*(3), 545-559.

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service?. *Bilingualism: Language and Cognition*, *13*(3), 359-371.

Brysbaert, M., & Ellis, A. W. (2016). Aphasia and age of acquisition: are early-learned words more resilient?. *Aphasiology*, *30*(11), 1240-1263.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977-990.

Brysbaert, M., Keuleers, E., & New, B. (2011). Assessing the usefulness of Google Books' word frequencies for psycholinguistic research on word processing. *Frontiers in Psychology, 2*, 1-8. https://doi.org/10.3389/fpsyg.2011.00027

Brysbaert, M., 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, 25*(1), 137–148. https://doi.org/10.1037/0096-1523.25.1.137.

Brysbaert, M., van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, *104*, 215–226. https://doi.org/10.1016/S0001-6918(00)00021-4

Caramazza, A., & Brones, I. (1980). Semantic classification by bilinguals. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *34*(1), 77-81. https://doi.org/10.1037/h0081016

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *The Quarterly Journal of Experimental Psychology*, *25*(1), 85-95.

Catling, J. C., Dent, K., & Williamson, S. (2008). Age of acquisition, not word frequency affects object recognition: Evidence from the effects of visual degradation. *Acta Psychologica*, *129*(1), 130-137.

Catling, J., Dent, K., Preece, E., & Johnston, R. (2013). Age-of-acquisition effects in novel picture naming: A laboratory analogue. *Quarterly Journal of Experimental Psychology*, *66*(9), 1756-1763.

Chen, H. C., & Leung, Y. S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 316–325. https://doi.org/10.1037/0278-7393.15.2.316

Chen, H. C., Vaid, J., Boas, D. A., & Bortfeld, H. (2011). Examining the phonological neighborhood density effect using near infrared spectroscopy. *Human Brain Mapping*, *32*(9), 1363-1370.

Chen, L., & Boland, J. E. (2008). Dominance and context effects on activation of alternative homophone meanings. *Memory & Cognition*, *36*(7), 1306-1323.

Chetail, F. (2015). Reconsidering the role of orthographic redundancy in visual word recognition. *Frontiers in psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00645

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review*, *3*(3), 149-210.

Clopper, C. G. (2021). Perception of dialect variation. *The handbook of speech perception,* 333-364.

Colbert-Getz, J., & Cook, A. E. (2013). Revisiting effects of contextual strength on the subordinate bias effect: Evidence from eye movements. *Memory & Cognition*, *41*(8), 1172-1184.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497-505.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. *Attention and Performance VI.* 535–555.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological review*, *108*(1), 204-256.

Coltheart, V., Laxon, V. J., & Keating, C. (1988). Effects of word imageability and age of acquisition on children's reading. *British Journal of Psychology*, *79*(1), 1-12.

Comesaña, M., Ferré, P., Romero, J., Guasch, M., Soares, A. P., & García-Chico, T. (2015). Facilitative effect of cognate words vanishes when reducing the orthographic overlap: The role of stimuli list composition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 614–635. https://doi.org/10.1037/xlm0000065

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.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1084-1096.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory & Cognition*, *2*(1), 130-138.

Conrad, M., Carreiras, M., Tamm, S., & Jacobs, A. M. (2009). Syllables and bigrams: orthographic redundancy and syllabic units affect visual word recognition at different processing levels. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 461-479.

Colorado, C. (2007). *Capitalizing on Similarities and Differences between Spanish and English.* Retrieved from: https://www.colorincolorado.org/article/capitalizing-similarities-and-differences-between-spanish-and-english#:~:text=Fortunately%20for%20Spanish%2Dspeaking%20English,a%20related%20word%20in%20Spanish.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *The Quarterly Journal of Experimental Psychology, 60*(8), 1072-1082.

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *The Quarterly Journal of Experimental Psychology*, *66*(5), 946-972.

Cortese, M. J., Khanna, M. M., & Hacker, S. (2010). Recognition memory for 2,578 monosyllabic words. *Memory, 18*(6), 595-609.

Cortese, M. J., McCarty, D. P., & Schock, J. (2015). A mega recognition memory study of 2897 disyllabic words. *The Quarterly Journal of Experimental Psychology, 68*(8), 1489-1501.

Costa, A., & Pickering, M. J. (2019). The role of learning on bilinguals' lexical architecture: Beyond separated vs. integrated lexicons. *Bilingualism: Language and Cognition*, *22*(4), 685-686.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology, 38*(3), 367-393.

Cuetos, F., Barbón, A., Urrutia, M., & Domínguez, A. (2009). Determining the time course of lexical frequency and age of acquisition using ERP. *Clinical Neurophysiology*, *120*(2), 285-294.

Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2012). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica*, *33*(2), 133-143.

Cuetos, F., Nosti, M., Gutiérrez, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica: Revista de Metodología y Psicología Experimental, 32*(2), 133–143.

Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson education.

Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika*, *9*(3), 185-197.

De Bot, K., Cox, A., Ralston, S., Schaufeli, A., & Weltens, B. (1995). Lexical processing in bilinguals. *Second Language Research, 11*(1), 1-19.

De Bruijn, E. R., Dijkstra, T., Chwilla, D. J., & Schriefers, H. J. (2001). Language context effects on interlingual homograph recognition: evidence from event-related potentials and response times in semantic priming. *Bilingualism: Language and Cognition*, *4*(2), 155-168.

De Groot, A. M. (1984). Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *The Quarterly Journal of Experimental Psychology*, *36*(2), 253-280.

De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in a bilingual and a monolingual task: Support for nonselective access to bilingual memory. *Quarterly Journal of Experimental Psychology*, *53*(2), 397-428.

De Groot, A. M., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of memory and language*, *30*(1), 90-123.

De Groot, A. M., Borgwaldt, S., Bos, M., & Van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language, 47*(1), 91-124.

Di Betta, A.M., Okurowska, E., & Morgan, J. L. (2015). Camping in the graveyard? Processing of interlingual Polish-English homographs. XII International Symposium of Psycholinguistics, Valencia, July.

Diener, E., & Biswas-Diener, R. (2016). The replication crisis in psychology.

Diependaele, K., Brysbaert, M., & Neri, P. (2012). How noisy is lexical decision?. *Frontiers in psychology, 3*(348).

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology, 66*(5), 843-863.

Dijkstra, A. F. J., & van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175-197.

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J.F. Kroll & A.M.B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*, *178*, 201.

Dijkstra, T., & van Heuven, W. J. (2012). Word recognition in the bilingual brain. *The handbook of the neuropsychology of language. Chichester: Blackwell Publishing*, 451-471.

Dijkstra, T., De Bruijn, E., Schriefers, H., & Ten Brinke, S. (2000a). More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism Language and Cognition*, *3*(1), 69-78.

Dijkstra, T., Grainger, J., & van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496-518.

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 ;anguage*, *62*(3), 284-301.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000b). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language*, *42*(4), 445-464.

Dijkstra, T., Van Hell, J. G., & Brenders, P. (2015). Sentence context effects in bilingual word recognition: Cognate status, sentence language, and semantic constraint. *Bilingualism: Language and Cognition*, *18*(4), 597-613.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, *1*(1), 51-66.

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657-679.

Dong, Y., Gui, S., & MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism*, *8*(3), 221-238.

Dopkins, S., Morris, R. K., & Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language*, *31*(4), 461-476.

Duffy, S. A., Henderson, J. M., & Morris, R. K. (1989). Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 791-801.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*(4), 429-446.

Durlik, J., Szewczyk, J., Muszyński, M., & Wodniecka, Z. (2016). Interference and inhibition in bilingual language comprehension: Evidence from Polish-English interlingual homographs. *PloS one*, *11*(3), e0151430.

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

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1103–1123. https://doi.org/10.1037/0278-7393.26.5.1103.

Ellis, A. W., Burani, C., Izura, C., Bromiley, A., & Venneri, A. (2006). Traces of vocabulary acquisition in the brain: Evidence from covert object naming. *NeuroImage*, *33*(3), 958-968.

Elston-Güttler, K. E., & Friederici, A. D. (2005). Native and L2 processing of homonyms in sentential context. *Journal of Memory and Language*, *52*(2), 256-283.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005b). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research*, *25*(1), 57-70.

Elston-Güttler, K. E., Paulmann, S., & Kotz, S. A. (2005a). Who's in control? Proficiency and L1 influence on L2 processing. *Journal of Cognitive Neuroscience*, *17*(10), 1593-1610

Ethnologue. (2021). *What are the top 200 most spoken languages?.* Retrieved from: https://www.ethnologue.com/guides/ethnologue200

Eurobarometer, S. (2006). Europeans and their Languages. *European Commission*.

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., ... & Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in Psychology*, *2*(306).

Ferreira, F., & Clifton Jr, C. (1986). The independence of syntactic processing. *Journal of memory and language*, *25*(3), 348-368.

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 555- 568. https://doi.org/10.1037/0278-7393.16.4.555

Fiebach, C. J., Friederici, A. D., Müller, K., Von Cramon, D. Y., & Hernandez, A. E. (2003). Distinct brain representations for early and late learned words. *NeuroImage*, *19*(4), 1627-1637.

Fitzpatrick, T., Playfoot, D., Wray, A., & Wright, M. J. (2015). Establishing the reliability of word association data for investigating individual and group differences. *Applied Linguistics*, *36*(1), 23-50.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing?. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276.

Font, N. (2001). *Rôle de la langue dans l'accès au lexique chez les bilingues: Influence de la proximité orthographique et sémantique interlangue sur la reconnaissance visuelle de mots* (Doctoral dissertation, Montpellier 3).

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 696-713. https://doi.org/10.1037/0278-7393.22.3.696

Foss, D. J. (1970). Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *9*(6), 699-706.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(3), 361–379. https://doi.org/10.1037/0096-1523.2.3.361

Gadsby, N., Arnott, W. L., & Copland, D. A. (2008). An investigation of working memory influences on lexical ambiguity resolution. *Neuropsychology*, *22*(2), 209 –216. https://doi.org/10.1037/0894-4105.22.2.209

Galimberti Jarman, B., Russell, R., Styles, C., Parker, S., & Huelskamp, C. (1994). The Oxford Spanish dictionary, Spanish-English/English-Spanish. Oxford (United Kingdom) Oxford Univ. Press

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevelhierarchical models* (Vol. 1). New York, NY, USA: Cambridge University Press.

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 305-315.

Gerhand, S., & Barry, C. (1999). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory & Cognition*, *27*(4), 592-602.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256 –281. https://doi.org/10.1037/0096-3445.113.2.256

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 430 –445. https://doi.org/10.1037/0278-7393.16.3.430

Ghyselinck, M., De Moor, W., & Brysbaert, M. (2000). Age-of-acquisition ratings for 2816 Dutch four- and five-letter nouns. *Psychologica Belgica*, *40*, 77–98.

Gilhooly, K. J., & Gilhooly, M. L. (1980). The validity of age-of-acquisition ratings. *British Journal of Psychology*, *71*(1), 105-110.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior research Methods & Instrumentation*, *12*(4), 395-427.

Glucksberg, S., Kreuz, R. J., & Rho, S. H. (1986). Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(3), 323 –335. https://doi.org/10.1037/0278-7393.12.3.323

Gottlob, L. R., Goldinger, S. D., Stone, G. O., & Van Orden, G. C. (1999). Reading homographs: Orthographic, phonologic, and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(2), 561 –574. https://doi.org/10.1037/0096-1523.25.2.561

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, *23*(1), 1-35.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review, 103*(3), 518 –565. https://doi.org/10.1037/0033-295X.103.3.518

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67-81.

Green, D. W. (2003). Neural basis of lexicon and grammar in L2 acquisition: The convergence hypothesis. In J. Benjamins (Ed) *The Lexicon–Syntax Interface in Second Language Acquisition* (pp. 197-218).

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual & Multicultural Development*, *6*(6), 467-477.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3-15.

Grosjean, F. (1992). Another view of bilingualism. In *Advances in psychology* (Vol. 83, pp. 51-62). North-Holland.

Grosjean, F. (1999). The bilingual's language modes. In J.L. Nicol (ed.) *One Mind, Two languages: Bilingual Language Processing* (preprint). Oxford: Blackwell.

Grojean, F. (2001). The Bilingual's language mode. In J.L. Nicol (ed.) *One Mind, Two Languages: Bilingual Language Processing.* Oxford: OUP.

Grosjean, F. (2010). *Bilingual*. Harvard university press.

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263-277.

Haugen, E. I. (1953). *The Norwegian language in America: The bilingual community* (Vol. 1). University of Pennsylvania Press.

Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch?. *Current Directions in Psychological Science, 10*(5), 164-168.

Heredia, R. R., & Cieślicka, A. B. (2014). Bilingual memory storage: Compound-coordinate and derivatives. In *Foundations of bilingual memory* (pp. 11-39). Springer, New York, NY.

Hicks, J. L., Franks, B. A., & Spitler, S. N. (2017). Prior task experience and comparable stimulus exposure nullify focal and nonfocal prospective memory retrieval differences. *The Quarterly Journal of Experimental Psychology*, *70*(10), 1997-2006.

Hill, H., Strube, M., Roesch-Ely, D., & Weisbrod, M. (2002). Automatic vs. controlled processes in semantic priming—differentiation by event-related potentials. *International Journal of Psychophysiology, 44*(3), 197-218.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1331 –1356. https://doi.org/10.1037/0096-1523.22.6.1331

Hino, Y., & Lupker, S. J. (2000). Effects of word frequency and spelling–to–sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 166 –183. https://doi.org/10.1037/0096-1523.26.1.166

Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition, 6*(2), 117-128.

Hoffmann, C. (2014). *Introduction to bilingualism*. Routledge.

Holcomb, P. J., Grainger, J., & O'rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, *14*(6), 938-950.

Holmes, S. J., Fitch, J. F., & Ellis, A. W. (2006). Age of acquisition affects object recognition and naming in patients with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 1010-1022.

Hornby, P. A. (1977). *Bilingualism: Psychological, social, and educational implications*. Academic Pr.

Hoversten, L. J., & Traxler, M. J. (2016). A time course analysis of interlingual homograph processing: Evidence from eye movements. *Bilingualism*, *19*(2), 347-360.

Hudson, P. T., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language*, *24*(1), 46-58.

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, *15*(2), 422-433.

Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica*, *23*(2).

Izura, C., & Ellis, A. W. (2004). Age of acquisition effects in translation judgement tasks. *Journal of Memory and Language, 50*(2), 165-181.

Izura, C., & Playfoot, D., (2012). A normative study of acronyms and acronym naming. *Behaviour Research Methods, 44*, 862-889.

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica*, *35*(1), 49-66.

Izura, C., Hernández-Muñoz, N., & Ellis, A. W. (2005). Category norms for 500 Spanish words in five semantic categories. *Behavior Research Methods, 37*, 385-397.

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language, 64*(1), 32-58.

Jackendoff, R., & Jackendoff, R. S. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press, USA.

Johnsrude, I. S., & Rodd, J. M. (2016). Factors that increase processing demands when listening to speech. In *Neurobiology of language* (pp. 491-502). Academic Press.

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual cognition*, *13*(7-8), 789-845.

Jouravlev, O., & Jared, D. (2014). Reading Russian-English homographs in sentence contexts: Evidence from ERPs. *Bilingualism*, *17*(1), 153-168.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*, 684–712. https://doi.org/10.1037/0033-2909.131.5.684

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition*, *13*(7-8), 846-863.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329 –354. https://doi.org/10.1037/0033-295X.87.4.329

Keatley, C. W., Spinks, J. A., & De Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, *22*(1), 70-84.

Kennet, J., McGuire, L., Willis, S. L., & Warner Schaie, K. (2000). Memorability functions in verbal memory: A longitudinal approach. *Experimental Aging Research, 26* (2), 121-137.

Kerkhofs, R., Dijkstra, T., Chwilla, D. J., & De Bruijn, E. R. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Research*, *1068*(1), 170-183.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44*(1), 287-304. doi: 10.3758/S13428- 011-0118-4

Kiefer, M., Weisbrod, M., Kern, I., Maier, S., & Spitzer, M. (1998). Right hemisphere activation during indirect semantic priming: Evidence from event-related potentials. *Brain and Language*, *64*(3), 377-408.

Kiran, S., & Tuchtenhagen, J. (2005). Imageability effects in normal Spanish–English bilingual adults and in aphasia: Evidence from naming to definition and semantic priming tasks. *Aphasiology, 19*(3-5), 315-327.

Klepousniotou, E. (2002). The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language*, *81*(1-3), 205-223.

Kolers, P. A. (1968). Bilingualism and information processing. *Scientific American*, *218*(3), 78-89.

Kotchoubey, B., & El-Khoury, S. (2014). Event-related potentials indicate context effect in reading ambiguous words. *Brain and Cognition*, *92,* 48-60.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 804 –823. https://doi.org/10.1037/0278-7393.20.4.804

Kroll, J. F., & De Groot, A. M. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A.M. De Groot and J. F. Kroll (eds) Tutorials in *Bilingualism: Psycholinguistic Perspectives* (pp. 169199). Mawhah, NJ: Lawrence Erlbaum.

Kroll, J. F., & De Groot, A. M. (2002). Lexical and conceptual memory in the bilingual. *W: Altmann GTM (red.), Psycholinguistics. Critical Concepts in Psychology*, *2*, 203-234.

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), The Oxford handbook of language production (pp. 165–181). Oxford University Press.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 92-107.

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., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism (Cambridge, England), 13*(3), 373-381. https://doi.org/10.1017/S136672891000009X

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown university press.

Kutateladze, M. (2015). Major differences between American and British English in business communication. *Journal in Humanities, 3*(2), 23-26.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211 –240. https://doi.org/10.1037/0033-295X.104.2.211

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2-3), 259-284.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford university press.

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language, 92*, 217-233.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*(2), 325-343.

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

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 12 –31. https://doi.org/10.1037/0278-7393.34.1.12

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59-67.

Levenshtein, V. I. (1965). Binary codes with correction for deletions and insertions of the symbol 1. *Problemy Peredachi Informatsii*, *1*(1), 12-25.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10*(8), 707–710.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, memory, and cognition*, *35*(2), 381- –390. https://doi.org/10.1037/a0014875

Liu, Y., Hao, M., Shu, H., Tan, L. H., & Weekes, B. S. (2008). Age-of-acquisition effects on oral reading in Chinese. *Psychonomic Bulletin & Review*, *15*(2), 344-350.

Love, T., & Swinney, D. (1996). Coreference processing and levels of analysis in object-relative constructions; demonstration of antecedent reactivation with the cross-modal priming paradigm. *Journal of Psycholinguistic Research*, *25*(1), 5-24.

Love, T., Maas, E., & Swinney, D. (2003). The influence of language exposure on lexical and syntactic language processing. *Experimental Psychology*, *50*(3), 204 –216. https://doi.org/10.1026/1617-3169.50.3.204

Lucas, M. (1999). Context effects in lexical access: A meta-analysis. *Memory & Cognition*, *27*(3), 385-398.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1-36.

Luce, P.A., Pisoni, D., & Goldinger, S. (1990) Similarity neighborhoods of spoken words. *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. 122–147.

Luk, G., Bialystok, E., Craik, F. I., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *Journal of Neuroscience*, *31*(46), 16808-16813.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203-208.

Luo, L., Craik, F. I., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: evidence from aging. *Psychology and Aging*, *28*(1), 28 –34. https://doi.org/10.1037/a0030875

Maciejewski, G., & Klepousniotou, E. (2016). Relative meaning frequencies for 100 homonyms: British eDom norms. *Journal of Open Psychology Data*, *4*.

Macizo, P., Bajo, T., & Martín, M. C. (2010). Inhibitory processes in bilingual language comprehension: Evidence from Spanish–English interlexical homographs. *Journal of Memory and Language*, *63*(2), 232-244.

Mackey, W. F. (1962). The description of bilingualism. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *7*(2), 51-85.

Marian, V., & Shook, A. (2012, September). The cognitive benefits of being bilingual. In Cerebrum: the Dana forum on brain science (Vol. 2012). Dana Foundation.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PloS One*, *7*(8), e4230.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, *244*(5417), 522-523.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, *189*(4198), 226-228.

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of discourse context as a determinant of the subordinate bias effect. *The Quarterly Journal of Experimental Psychology Section A*, *52*(4), 813-839.

Martín, M. C., Macizo, P., & Bajo, T. (2010). Time course of inhibitory processes in bilingual language processing. *British Journal of Psychology*, *101*(4), 679-693.

Massaro, D. W., & Cohen, M. M. (1994). Visual, orthographic, phonological, and lexical influences in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1107 –1128. https://doi.org/10.1037/0096-

Mazur, B. W. (2011). *Colloquial Polish: the complete course for beginners* (Vol. 1). Routledge.

McClelland, J. L. (1987). *The case for interactionism in language processing*. CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF PSYCHOLOGY.

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*, *88*(5), 375 –407. https://doi.org/10.1037/0033-295X.88.5.375

Mellet, E., Tzourio, N., Denis, M., & Mazoyer, B. (1998). Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *Neuroreport*, *9*(5), 803-808.

Menenti, L., & Burani, C. (2007). What causes the effect of age of acquisition in lexical processing?. *Quarterly Journal of Experimental Psychology, 60*(5), 652-660.

Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition*, *17*(1), 89-117.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*(1), 25-40.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). In PMA Rabbitt. *Attention and Performance V,* 98-118.

Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. Sage.

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling--sound consistency in word naming?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 183–206. https://doi.org/10.1037/0278-7393.28.1.183

Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language, 63*, 506–525. https://doi.org/10.1016/j.jml.2010.08.003

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they?. *Journal of Experimental Psychology: General*, *118*(1), 43-71.

Montefinese, M., Vinson, D., Vigliocco, G., & Ambrosini, E. (2019). Italian age of acquisition norms for a large set of words (ItAoA). *Frontiers in Psychology, 10*(278).

Moradi, H. (2014). An investigation through different types of bilinguals and bilingualism. *International Journal of Humanities & Social Science Studies*, *1*(2), 147-154.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 116 –133. https://doi.org/10.1037/0278-7393.21.1.116

Mortensen, L., Berntsen, D., & Bohn, O. S. (2015). Retrieval of bilingual autobiographical memories: Effects of cue language and cue imageability. *Memory*, *23*(2), 138-156.

Mulatti, C., Besner, D., & Job, R. (2003). Reading aloud: Evidence for a phonologically based neighborhood effect. In *44th Annual Meeting of the Psychonomic Society, Vancouver, BC, Canada*.

Mulatti, C., Reynolds, M. G., & Besner, D. (2006). Neighborhood effects in reading aloud: New findings and new challenges for computational models. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 799–810. https://doi.org/10.1037/0096-1523.32.4.799

Mulder, K., Dijkstra, T., Schreuder, R., & Baayen, H. R. (2014). Effects of primary and secondary morphological family size in monolingual and bilingual word processing. *Journal of Memory and Language*, *72*, 59-84.

Mulder, K., Schreuder, R., & Dijkstra, T. (2013). Morphological family size effects in L1 and L2 processing: an electrophysiological study. *Language and Cognitive Processes*, *28*(7), 1004-1035.

Mulder, K., van Heuven, W. J., & Dijkstra, T. (2018). Revisiting the neighborhood: How L2 proficiency and neighborhood manipulation affect bilingual processing. *Frontiers in Psychology*, *9*(1860). https://doi.org/10.3389/fpsyg.2018.01860

Nas, G. (1983). Visual word recognition in bilinguals: Evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior*, *22*(5), 526-534.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 15(6), 1003 –1019. https://doi.org/10.1037/0278-7393.15.6.1003

New, B. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*(1), 45-52.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*(4), 661-677.

New, B., Ferrand, L., Pallier, C., and Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: new evidence from the English Lexicon Project. *Psychon. Bull. Rev.* 13, 45–52.

Oden, G. C., & Spira, J. L. (1983). Influence of context on the activation and selection of ambiguous word senses. *The Quarterly Journal of Experimental Psychology*, *35*(1), 51-64.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, *9*(3), 225-236.

Owsowitz, S. E. (1943). *The effects of word familiarity and letter structure familiarity on the perception of words* (No. P-2820). RAND CORP SANTA MONICA CA.

Owsowitz, S. E. (1963). *The Effects of Word Familiarity and Letter Structure Familiarity on the Perception of Words.* Santa Monica, CA: Rand Corporation Publications.

Paivio, A. (1963). Learning of adjective-noun paired associates as a function of adjective-noun word order and noun abstractness. *Canadian Journal of Psychology/Revue canadienne de psychologie, 17*(4), 370 –379. https://doi.org/10.1037/h0083277

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, *4*(1), 32-38.

Paivio, A. (1967). Paired-associate learning and free recall of nouns as a function of concreteness, specificity, imagery, and meaningfulness. *Psychological Reports*, *20*(1), 239-245.

Paivio, A. (1971). Imagery and language. In *Imagery* (pp. 7-32). Academic Press.

Paivio, A. (2013). *Imagery and verbal processes*. Psychology Press.

Paivio, A. (2014). Bilingual dual coding theory and memory. In *Foundations of bilingual memory* (pp. 41-62). Springer, New York, NY.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1p2), 1.

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.

Perea, M., & Rosa, E. (2002). Does “whole-word shape” play a role in visual word recognition?. *Perception & Psychophysics, 64*(5), 785-794.

Pérez, M. A. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology*, *60*(1), 32-42.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383.

Peter, J. P. (1979). Reliability: A review of psychometric basics and recent marketing practices. *Journal of Marketing Research*, *16*(1), 6-17.

Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects?. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *53*(4), 323–334. https://doi.org/10.1037/h0087320

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 787–796. https://doi.org/10.1037/a0035583

Plaut, D. C., & Shallice, T. (1993). Perseverative and semantic influences on visual object naming errors in optic aphasia: A connectionist account. *Journal of Cognitive Neuroscience, 5*(1), 89-117.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.

Playfoot, D., Billington, J., & Tree, J. J. (2018). Reading and visual word recognition ability in semantic dementia is not predicted by semantic performance. *Neuropsychologia, 111*, 292-306.

Poncelet, M., Majerus, S., Raman, I., Warginaire, S., & Weekes, B. S. (2007). Naming actions and objects in bilingual aphasia: A multiple case study. *Brain and Language, 103*(1-2), 158-159.

Poort, E. D., & Rodd, J. M. (2017). The cognate facilitation effect in bilingual lexical decision is influenced by stimulus list composition. *Acta Psychologica, 180*, 52-63.

Poort, E. D., & Rodd, J. M. (2018). Does recent experience with identical cognates and interlingual homographs in one's native language affect processing of these words in one's second language. *Preregistration. http://www. osf. io/y6phs*.

Poort, E. D., & Rodd, J. M. (2019). Towards a distributed connectionist account of cognates and interlingual homographs: Evidence from semantic relatedness tasks. *PeerJ, 7, e6725.*

Poort, E. D., Warren, J. E., & Rodd, J. M. (2016). Recent experience with cognates and interlingual homographs in one language affects subsequent processing in another language. *Bilingualism: Language and Cognition*, *19*(1), 206-212.

Potter, M. C., So, K. F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 23-38.

Raman, I. (2006). On the age-of-acquisition effects in word naming and orthographic transparency: Mapping specific or universal?. *Visual Cognition*, *13*(7-8), 1044-1053.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. (Conference Paper).

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191-201.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 779-790.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language*, *33*(4), 527-544.

Rice, G. A., & Robinson, D. O. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory & Cognition*, *3*(5), 513-518.

Richards, R. M., & Ellis, A. W. (2009). Identity, gender, and the role of age of acquisition in face processing. *Psicológica*, *30*(2), 155-179.

Richardson, J. T. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology*, *67*(3), 315-325.

Ritchie, W. C., & Bhatia, T. K. (2013). *The Handbook of Bilingualism and Multilingualism*. Blackwell Publishing.

Robson, S. (2002). From Malay to Indonesian: The genesis of a national language. *Working Papers (Monash University. Centre of Southeast Asian Studies)*, iii.

Rodd, J. (2017). *Lexical ambiguity*. Oxford handbook of psycholinguistics.

Rodd, J. M., Cai, Z. G., Betts, H. N., Hanby, B., Hutchinson, C., & Adler, A. (2016). The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments. *Journal of Memory and Language*, *87*, 16-37.

Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, *68*(2), 180-198.

Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, *68*(2), 180-198.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*(2), 245-266.

Rofes, A., Zakariás, L., Ceder, K., Lind, M., Johansson, M. B., De Aguiar, V., & Sacristán, C. H. (2018). Imageability ratings across languages. *Behavior Research Methods*, *50*(3), 1187-1197.

Rogers, V., Playfoot, D., & Milton, J. (2018). 7. The Second Language Lexicon. In *Mind Matters in SLA* (pp. 117-136). Multilingual Matters.

Romaine, S. (1989). The role of children in linguistic change. *Breivik, LE & Jahr, EH eds*, 199-226.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 487-494.

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review*, *1*(2), 99.

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *Neuroimage*, *27*(1), 188-200.

Sáchez-Casas, R. M., García-Albea, J. E., & Davis, C. W. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology*, *4*(4), 293-310.

Salamoura, A., & Williams, J. N. (2007). Processing verb argument structure across languages: Evidence for shared representations in the bilingual lexicon. *Applied Psycholinguistics, 28*(4), 627-660.

Schallice, T., & Burgess, P. (1996). The domain of supervisory processes and temporal organization of behavior. *Philosophical Transactions of the Royal Society of London B*, *351*, 1405-1412.

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, *15*(1), 157-166.

Schmalz, X., & Mulatti, C. (2017). Busting a myth with the Bayes Factor. *The Mental Lexicon, 12*(2), 263-282.

Schulpen, B., Dijkstra, A., & Schriefers, H. J. (2003). L2 proficiency and task effects on interlingual homograph recognition by dutch-english bilinguals. *Manuscript, University of Nijmegen*.

Schvaneveldt, R. W., & Meyer, D. E. (1973). *Retrieval and comparison processes in semantic memory* (pp. 395–409). Attention and performance IV. New York, NY: Academic Pres.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27*(5), 499-520.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of memory and language*, *55*(2), 197-212.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*(4), 876-900.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Center for the Study of Reading Technical Report; no. 240*.

Sereno, S. C., Pacht, J. M., & Rayner, K. (1992). The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science, 3*(5), 296-301.

Shaw, M. E., Hawley, G. G., & Wright, J. M. (1967). *Scales for the measurement of attitudes.* New York: McGraw-Hill.

Shin, S. J. (2017). *Bilingualism in schools and society: Language, identity, and policy*. Routledge.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological science*, *6*(1), 45-49.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual review of psychology, 69,* 487-510.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior*, *20*(1), 120-136.

Simpson, G. B. (1994). Context and the processing of ambiguous words. *Handbook of Psycholinguistics*, *22*, 359-374.

Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(1), 28-39.

Skutnabb-Kangas, T. (1981). *Bilingualism or not: The education of minorities* (Vol. 7). Multilingual Matters.

Smith, C., Bermejo Marcos, M., & Chang-Rodríguez, E. (1986). Collins Spanish-English, English-Spanish dictionary. Collins.

Stewart, N., & Ellis, A. W. (2008). Order of acquisition in learning perceptual categories: A laboratory analogue of the age-of-acquisition effect?. *Psychonomic Bulletin & Review, 15*(1), 70-74.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29,* 41–78. doi: 10.1207/s15516709cog2901_3

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1140 –1154. https://doi.org/10.1037/0278-7393.21.5.1140

Stroebe, W. (2016). Are most published social psychological findings false?. *Journal of Experimental Social Psychology, 66*, 134-144.

Stroebe, W. (2019). What Can We Learn from Many Labs Replications?. *Basic and Applied Social Psychology*, *41*(2), 91-103.

Struys, E., Woumans, E., Nour, S., Kepinska, O., & Van den Noort, M. (2019). A domain-general monitoring account of language switching in recognition tasks: Evidence for adaptive control. *Bilingualism: Language and Cognition, 22*(3), 606-623.

Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, 387-422.

Surrain, S. H., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*.

Swinney, D. A. (1979). Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*(6), 645-659.

Szmigiera, M. (2021). *The most spoken languages worldwide 2021*. Retrieved from: https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5, pp. 481-498). Boston, MA: Pearson.

Tabors, P. O., & Snow, C. E. (2001). Young bilingual children and early literacy development. *Handbook of Early Literacy Research, 1*, 159-178.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 2*(3), 324-340.

Tabossi, P. (1989). What's in a context?. In *Resolving semantic ambiguity* (pp. 25-39). Springer, New York, NY.

Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language, 32*(3), 359-372.

Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition, 2*(1), 45-58.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of verbal learning and verbal behavior, 18*(4), 427-440.

Thorndike, R. L. (1973). Reading as reasoning. *Reading Research Quarterly*, 135-147.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1412–1431. https://doi.org/10.1037/a0024492

Tottie, G. (2009). How different are American and British English grammar? And how are they different. *One Language, Two Grammars*, 341-363.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35*(3), 454-475.

Twilley, L. C., & Dixon, P. (2000). Meaning resolution processes for words: A parallel independent model. *Psychonomic Bulletin & Review*, *7*(1), 49-82.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, *22*(1), 111-126.

Valdés, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Ablex Publishing.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64*(1), 88-107.

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology, 3*, 174.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? *Cognate effects in a sentence context. Psychological science, 20*(8), 923-927.

Van Hell, J. G., & De Groot, A. M. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128*(3), 431-451.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*(4), 780-789.

van Heuven, W. J., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews*, *64*(1), 104-122.

van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*(3), 458-483.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190.

Van Kesteren, R., Dijkstra, T., & De Smedt, K. (2012). Markedness effects in Norwegian–English bilinguals: Task-dependent use of language-specific letters and bigrams. *Quarterly Journal of Experimental Psychology*, *65*(11), 2129-2154.

Vanlangendonck, F., Peeters, D., Rueschemeyer, S. A., & Dijkstra, T. (2020). Mixing the stimulus list in bilingual lexical decision turns cognate facilitation effects into mirrored inhibition effects. *Bilingualism: Language and Cognition*, *23*(4), 836-844.

Vitello, S., & Rodd, J. M. (2015). Resolving semantic ambiguities in sentences: Cognitive processes and brain mechanisms. *Language and Linguistics Compass*, *9*(10), 391-405.

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. (2002). Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism*, *5*(1), 1-23.

Vu, H., Kellas, G., & Paul, S. T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & Cognition*, *26*(5), 979-1001.

Weekes, B. S., Chan, A. H., & Tan, L. H. (2008). Effects of age of acquisition on brain activation during Chinese character recognition. *Neuropsychologia*, *46*(7), 2086-2090.

Weinreich, U. (1953). Languages in contact, linguistic circle of New York. *Publication*, *1*, 64-65.

Weinreich, U. (1968). *Languages in contact.* The Hague: Mouton.

Westbury, C., & Buchanan, L. (2002). The probability of the least likely non-length-controlled bigram affects lexical decision reaction times. *Brain and Language*, *81*(1-3), 66-78.

Wilson, M. A., Cuetos, F., Davies, R., & Burani, C. (2013). Revisiting age-of-acquisition effects in Spanish visual word recognition: The role of item imageability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6).

Wilson, M. A., Ellis, A. W., & Burani, C. (2012). Age-of-acquisition affects word naming in Italian only when stress is irregular. *Acta Psychologica*, *139*(3), 417-424.

Wray, A. (2002). Transition to Language. Oxford University Press UK.

Wu, Y. J., & Thierry, G. (2010). Investigating bilingual processing: The neglected role of language processing contexts. *Frontiers in psychology*, *1*, 178.

Yang, E. (2017). Bilinguals' working memory (WM) advantage and their dual language practices. *Brain sciences*, *7*(7), 86. https://doi.org/10.3390/brainsci7070086

Yap, M. J., Balota, D. A., Tse, C. S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 495 –513. https://doi.org/10.1037/0278-7393.34.3.495

Yarmey, A. D., & O'Neill, B. J. (1969). SR and RS paired-associate learning as a function of concreteness, imagery, specificity, and association value. *The Journal of Psychology*, *71*(1), 95-109.

Yates, M. (2005). Phonological neighbors speed visual word processing: Evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1385 –1397. https://doi.org/10.1037/0278-7393.31.6.1385

Yates, M., Locker, L., & Simpson, G. B. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review*, *11*(3), 452-457.

Zeno, S., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Touchstone Applied Science Associates.

Zipf, G. K. (1936). *The Psycho-Biology of Language*. Routledge, London.

Zipf, G. K. (1949). Human behavior and the principle of least effort: an introd. to human ecology. Cambridge, Mass.

Appendices (see separate volume)